

STIC-Biotech/ChemLib

89/22

From: Slobodyansky, Elizabeth
Sent: Friday, March 14, 2003 3:56 PM
To: STIC-Biotech/ChemLib
Subject: 09/744,125

Please search for case 09/744,125:

EL

SEQ ID NOS: 1 and 2 against commercial and interference databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner

Art Unit 1652
CM1 10D11
703-306-3222

mail box 10D01

Mary Jane Ruhl
Tech. Info. Specialist, STIC
TC-1600
CM-1, Room 6A-06
Phone: 605-1155

RECEIVED
MAR 14 2003
STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 3/8
Date Completed: 3/31
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

This Page Blank (uspto)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 27, 2003, 05:22:50 ; Search time 446 Seconds

(without alignments)
7841.600 Million cell updates/sec

Title: US-09-744-125A-1

Perfect score: 1553
Sequence: 1 gaattcgccagcaggtcg999.....aaaaaaaaaacctcgag 1553

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	100.0	1553	21	AAZ58953 Human MIF1 protein
2	1535.8	98.9	1914	21	AAZ58958 Human MIF1 protein
3	414.4	26.7	10579	22	ABA20014 Human nervous syst
4	414.4	26.7	10579	22	AAK70045 Human immune/haema
5	414.4	26.7	28995	22	ABA20015 Human nervous syst
6	414.4	26.7	28995	22	AAK70046 Human immune/haema
7	414.4	26.7	28995	22	AAK79667 Human immune/haema
8	414.4	26.7	28995	22	AAK85213 Human immune/haema
9	351.4	22.6	1905	23	ABL02115 Drosophila melanog

C	10	285.6	18.4	4027	23	ABL02114	Drosophila melanog
	11	159.4	10.3	8204	22	AAK79966	Human immune/haema
	12	159.4	10.3	8204	22	AAK85212	Human immune/haema
	13	121	7.8	144	16	AA723953	Human gene signatu
	14	71.6	4.6	457	22	AAK87783	Human digestive sy
	15	71.6	4.6	457	22	AAK81700	Human liver asocci
	16	71.6	4.6	457	22	ABN90055	Human liver antige
	17	60.6	3.9	20974	23	AAK92595	DNA encoding novel
	18	60.2	3.9	65	24	ABN52890	Mouse spliced tran
	19	60	3.9	60	24	ABN59396	Human spliced tran
	20	58	3.7	2108	24	ABU67774	Oesophagus cancer
	21	50.2	3.2	60	24	ABN59283	Human spliced tran
	22	50.2	3.2	60	24	ABN59397	Human spliced tran
	23	49.4	3.2	51	22	AAJ30900	Human SNP oligonuc
	24	49.4	3.2	799	19	AAV55831	Nucleotide sequenc
	25	49.4	3.2	1925	20	AAK90924	Epsstein-Barr virus
	26	49.4	3.2	1926	21	AAK50254	Epsstein-Barr virus
	27	49.4	3.2	1926	22	AAK82902	EBV tethering prot
	28	49.4	3.2	2580	21	AAK75454	Nucleotide sequenc
	29	49.4	3.2	2580	24	AAK64275	Epsstein-Barr virus
	30	49.4	3.2	5452	20	AAK90923	Anti-sense strand
	31	49.4	3.2	8705	20	AAK23378	Vector pShuttle DN
	32	49.4	3.2	9600	19	AAK21683	Vector plasmid PCM
	33	49.4	3.2	10380	20	AAK22248	Nucleotide sequenc
	34	49.4	3.2	10596	14	AAK051731	Plasmid pCisEBON f
	35	49.4	3.2	10596	17	AAK40348	Plasmid pCisEBON f
	36	49.4	3.2	10596	20	AAK15650	Nucleotide sequenc
	37	49.4	3.2	16080	21	AAK59553	DNA clone PCR Cl.
	38	48.6	3.1	3489	21	AAK10290	Kaposi's sarcoma-a
	39	48.6	3.1	3489	22	AAK82901	Nucleotide sequenc
	40	48.6	3.1	3489	24	ABK93487	Kaposi's sarcoma-a
	41	48.6	3.1	32207	24	AAK73805	KSHV LTR DNA (nucl
	42	48.6	3.1	13207	19	AAV19941	KSHV long unique c
	43	47.6	3.0	16442	18	AAK83006	Partial mouse WRN
	44	47.2	3.0	6522	24	ABU33051	Human immune syste
	45	47	3.0	5120	22	AAK84677	DNA sequence of hu

ALIGNMENTS

RESULT 1	AAZ58953	AAZ58953 standard; cDNA; 1553 BP.
XX	AAZ58953;	
AC	08-MAY-2000	(first entry)
XX		
DT		
XX		
DE		Human MIF1 protein encoding cDNA (plasmid pCM480).
XX		
KW	MEK kinase; MEKK interacting forkhead associated protein; MIF1, MEKK;	
KW	FHA protein; forkhead associated protein; tumour; angiogenesis; human;	
KW	psoriasis; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;	
KW	vasotropic; neuroprotective; antiarthritic; antiviral; ss.	
OS	Homo sapiens.	
XX		
FX	Key	Location/Qualifiers
FT	mat_peptide	2..1174
FT		/*tag= a
XX		
FX	WO200005362-A1.	
XX		
PD	03-FEB-2000.	
XX		
PF	21-JUL-1999;	99WO-EP05142.
XX		
PR	21-JUL-1998;	98US-0093590.
XX		
PA	(RHON) RHONE-POULENC RORER SA.	
XX		
PI	Marcireau C, Multon M, Polard-Houeset V;	

XX MPI: 2000-195102/17.
DR P-PSDB; AAU77554.
XX
PT New MEK kinase interacting forkhead associated protein (MIF1) useful to
PT treat or diagnose, e.g. inflammation and tumors, and to identify its
PT specific modulators, to regulate MEK kinase activity -
XX
PS Claim 3; Page 64-67; 78pp; English.
XX
XX The invention provides MEK kinase (MEKK) interacting forkhead associated
CC (FHA) protein (MIF1). MIF1 is useful for screening for specific
CC modulators (potential therapeutic agents) and to reduce MEK activity in
CC cells. Antibodies specific to MIF1 are useful as diagnostic immunoassay
CC reagents to detect expression of MIF1, for purification of MIF1 and as
CC therapeutic (ant)agonists. The MIF1 nucleic acids are useful for:
CC recombinant production of MIF1, either in cultured cells or in vivo (gene
CC therapy); as source of probes and primers for detecting or quantifying
CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source
CC of therapeutic antisense sequences (used to increase MEK activity in
CC cells); and to identify inhibitors of MIF1. Regulation of MEK activity
CC via MIF1 regulation is useful for treatment of inflammation, asthma,
CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic
CC syndrome, neurodegeneration, tumours, angiogenesis, rheumatoid arthritis,
CC psoriasis or persistent viral infections. The present sequence represents
CC a cDNA (plasmid pCM480) encoding a MIF1 protein.

XX Sequence 1553 BP; 367 A; 457 C; 423 G; 306 T; 0 other;

Query Match 100.0%; Score 1553; DB 21; Length 1553;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGACGACGAGTGGGGTGGACACGAGCGCTGTTCGGGGAGTGAACCTCTCCAG 60
DB 1 GAATTCGACGACGAGTGGGGTGGACACGAGCGCTGTTCGGGGAGTGAACCTCTCCAG 60
QY 61 TGAAGAAAAGAAAGTATCCAAAGCCCCCAGACATCTCTGTGCCACCCAGCCCGCCAGC 120
DB 61 TGAAGAAAAGAAAGTATCCAAAGCCCCCAGACATCTCTGTGCCACCCAGCCCGCCAGC 120
QY 121 CCTGGAATCACCAGACGCTGTGAAGAGTAAACAGCACTTCAGGTGACCAAGATCT 180
DB 121 CCTGGAATCACCAGACGCTGTGAAGAGTAAACAGCACTTCAGGTGACCAAGATCT 180
QY 121 CCTGGAATCACCAGACGCTGTGAAGAGTAAACAGCACTTCAGGTGACCAAGATCT 180
DB 121 CCTGGAATCACCAGACGCTGTGAAGAGTAAACAGCACTTCAGGTGACCAAGATCT 180
QY 181 GGGCGCGTGAAGCTGCAATGACCTCTCTCATTAATGCTGTGTGCAACCAAGA 240
DB 181 GGGCGCGTGAAGCTGCAATGACCTCTCTCATTAATGCTGTGTGCAACCAAGA 240
QY 241 CCTGACCTCCGTCACCTGCGGTGAATTCAGCTGCGCTTCACCTTCGGAGGTGCA 300
DB 241 CCTGACCTCCGTCACCTGCGGTGAATTCAGCTGCGCTTCACCTTCGGAGGTGCA 300
QY 301 GGAGCGTTGTAGCGCCTCTCTCATGATCTCTCAAGTTGGCTGTGAGGCCAT 360
DB 301 GGAGCGTTGTAGCGCCTCTCTCATGATCTCTCAAGTTGGCTGTGAGGCCAT 360
QY 361 GAGGAGCTGCAACCAAGAGCTATTGACAGCCATCCAGAGCAAGGCCCTGTTAGCAAGGC 420
DB 361 GAGGAGCTGCAACCAAGAGCTATTGACAGCCATCCAGAGCAAGGCCCTGTTAGCAAGGC 420
QY 421 TGAAGAGAGCTGCTGAGCAAAAGTGGATTCAGACAGCCAGCCCACTTGGAGACTTTCCA 480
DB 421 TGAAGAGAGCTGCTGAGCAAAAGTGGATTCAGACAGCCAGCCCACTTGGAGACTTTCCA 480
QY 481 GGACCTGCTGACAGACACCTCTGATGCTGACCGCCGTAACCGCAAGGCCCTTGCA 540
DB 481 GGACCTGCTGACAGACACCTCTGATGCTGACCGCCGTAACCGCAAGGCCCTTGCA 540
QY 541 GGCCCACTGAGCACTGATGAAGAGTATTAACCTGCTGAGAGACAGACAGTGCAGCGCT 600
DB 541 GGCCCACTGAGCACTGATGAAGAGTATTAACCTGCTGAGAGACAGACAGTGCAGCGCT 600

QY 601 GCCCAAGGGGACCAAGTGTCTGAACTTCTGTGATGACAGAGCACTGATGATGACATGA 660
DB 601 GCCCAAGGGGACCAAGTGTCTGAACTTCTGTGATGACAGAGCACTGATGATGACATGA 660
QY 661 GCTCAAGGACATGCGAGATGAGGTCTTGAACATGAGCTGATGTGTGCTGACCGGCCCA 720
DB 661 GCTCAAGGACATGCGAGATGAGGTCTTGAACATGAGCTGATGTGTGCTGACCGGCCCA 720
QY 721 GAAGGAGAGATTTGGCGAGCTGGGAACAGAACTGATTAAGTGGGAGGTCTGATGAGCAG 780
DB 721 GAAGGAGAGATTTGGCGAGCTGGGAACAGAACTGATTAAGTGGGAGGTCTGATGAGCAG 780
QY 781 CATCAAGGACATGAGCTCTCCGAGCTTGCACACAGCACTAGGAGTGTGCGGGGCG 840
DB 781 CATCAAGGACATGAGCTCTCCGAGCTTGCACACAGCACTAGGAGTGTGCGGGGCG 840
QY 841 CATGTCGCGTACCTGATGCGTGTGAGATCAACCTTGGGAGAGCAACAGAGATGA 900
DB 841 CATGTCGCGTACCTGATGCGTGTGAGATCAACCTTGGGAGAGCAACAGAGATGA 900
QY 901 CCAAGTTGATGAGACCTGTCTTGGAGAGTCCGGGCTGGAAAGATATCCGGAAACAGAG 960
DB 901 CCAAGTTGATGAGACCTGTCTTGGAGAGTCCGGGCTGGAAAGATATCCGGAAACAGAG 960
QY 961 TGTCAATCAAGCTGAAGAACCAACGAGTATTTCTTCAATGCGAATGAGGTGCAAGCCCAT 1020
DB 961 TGTCAATCAAGCTGAAGAACCAACGAGTATTTCTTCAATGCGAATGAGGTGCAAGCCCAT 1020
QY 1021 CTACATCGATGAGACGCGCGGTGCTGTGAGCTCAAAATGCGGCTCAGCAACACTCTGT 1080
DB 1021 CTACATCGATGAGACGCGCGGTGCTGTGAGCTCAAAATGCGGCTCAGCAACACTCTGT 1080
QY 1081 GGTGAGATTCGACACCTGCTCTTCTTCTTCAACAGAGACCTCATGCCCCCAT 1140
DB 1081 GGTGAGATTCGACACCTGCTCTTCTTCTTCAACAGAGACCTCATGCCCCCAT 1140
QY 1141 CAGGCTGAGGCTGCGCAAGATACACACAGTAGAAGATGAGGAGAGACTGTGGGCC 1200
DB 1141 CAGGCTGAGGCTGCGCAAGATACACACAGTAGAAGATGAGGAGAGACTGTGGGCC 1200
QY 1201 TCTCGGCTGTTTTCCCTGACACTCCAGCCCTTGAAGCTGGGAATCAGGCTCTGGA 1260
DB 1201 TCTCGGCTGTTTTCCCTGACACTCCAGCCCTTGAAGCTGGGAATCAGGCTCTGGA 1260
QY 1261 AAAAAGCTGGGAGTGGAGAGCTCACTGCGGGCAATTGATTTGAGAGAGAT 1320
DB 1261 AAAAAGCTGGGAGTGGAGAGCTCACTGCGGGCAATTGATTTGAGAGAGAT 1320
QY 1321 AGGCTGGGCTTTGTGAAGCCAGAGAGCTGAGAACTCAGGCTTCCCTAGATCCAGAG 1380
DB 1321 AGGCTGGGCTTTGTGAAGCCAGAGAGCTGAGAACTCAGGCTTCCCTAGATCCAGAG 1380
QY 1381 CCCCTCCCATCTTCTCTCTTAAACCAACCTTACCCCTTCTACCCCTCATTTGCC 1440
DB 1381 CCCCTCCCATCTTCTCTCTTAAACCAACCTTACCCCTTCTACCCCTCATTTGCC 1440
QY 1441 ACCTTCACTCTCTGTCTTCAAGCTGATTAAGCTGAGACTCTTCTTTATTTGTTTTCTT 1500
DB 1441 ACCTTCACTCTCTGTCTTCAAGCTGATTAAGCTGAGACTCTTCTTTATTTGTTTTCTT 1500
QY 1501 TGTAAATATAAAGCAACAGGTTCCAAAGTAAATAAATAAATAAATAAATAAATAAATAA 1553
DB 1501 TGTAAATATAAAGCAACAGGTTCCAAAGTAAATAAATAAATAAATAAATAAATAAATAA 1553

RESULT 2

AA258958
ID AA258958 standard; cDNA; 1914 BP.

AA258958;

08-MAY-2000 (first entry)

XX

DE Human MIF1 protein encoding cDNA (plasmid pCM577).
XX MEK kinase; MEKK interacting forkhead associated protein; MIF1; MEKK;
XX FHA protein; forkhead associated protein; tumour; angiogenesis; human;
KM peptidase; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;
KM vasotrophic; neuroprotective; antiarthritic; antiviral; ss.
XX
OS Homo sapiens.
FH Key location/Qualifiers
FT CDS 147..1535
FT /*tag= a
XX
XX MO200005362-A1.
XX
XX 03-FEB-2000.
XX
XX 21-JUL-1999; 99WO-EP05142.
XX
XX 21-JUL-1998; 98US-0093590.
XX
XX (RHON) RHONE-POULENC RORER SA.
XX
XX Marcireau C, Multon M, Polard-Housses V;
XX WPI: 2000-195102/17.
XX P-PSDB; AAY77555.
XX
XX New MEK kinase interacting forkhead associated protein (MIF1) useful to
PT treat or diagnose, e.g. inflammation and tumours, and to identify its
PT specific modulators, to regulate MEK kinase activity -
XX
XX
XX Claim 5; Page 70-73; 78pp; English.
XX
XX The invention provides MEK kinase (MEKK) interacting forkhead associated
CC (FHA) protein (MIF1). MIF1 is useful for screening for specific
CC modulators (potential therapeutic agents) and to reduce MEKK activity in
CC cells. Antibodies specific to MIF1 are useful as diagnostic immunosay
CC reagents to detect expression of MIF1, for purification of MIF1 and as
CC therapeutic (antagonists). The MIF1 nucleic acids are useful for:
CC recombinant production of MIF1, either in cultured cells or in vivo (gene
CC therapy); as source of probes and primers for detecting or quantifying
CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source
CC of therapeutic antisense sequences (used to increase MEKK activity in
CC cells); and to identify inhibitors of MIF1. Regulation of MEKK activity
CC via MIF1 regulation is useful for treatment of inflammation, asthma,
CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic
CC syndrome, neurodegeneration, tumours, angiogenesis, rheumatoid arthritis,
CC psoriasis or persistent viral infections. The present sequence represents
CC a CDNA (plasmid pCM577) encoding a MIF1 protein.
XX
XX
SQ Sequence 1914 BP; 449 A; 549 C; 534 G; 382 T; 0 other;
Query Match 98.9%; Score 1535.8; DB 21; Length 1914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1537; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 15 GTGGGGTGAACACAGGGCGCTTTGGGGAGTGAACCTCTCTCCAGTGAAGAAAGG 74
Db 376 GTGGGGTGAACACAGGGCGCTTTGGGGAGTGAACCTCTCTCCAGTGAAGAAAGG 435
QY 75 TATCCAAAGCCCCCGACACTCTGTGCAACCCAGCCAGCCCCCGCTGAGACACCA 134
Db 436 TATCCAAAGCCCCCGACACTCTGTGCAACCCAGCCAGCCCCCGCTGAGACACCA 495
QY 135 AGCGTGTGAAGAGTAACAGCACTTCAGGTGACCAAGATCTGGGCGCTGGAAGC 194
Db 496 AGCGTGTGAAGAGTAACAGCACTTCAGGTGACCAAGATCTGGGCGCTGGAAGC 555
QY 195 CTGCAATGACCTCTCTGCTCATTAATGCTGTGTGCAAGACCAAGACCTGACCTCCGTC 254
Db 556 CTGCAATGACCTCTCTGCTCATTAATGCTGTGTGCAAGACCAAGACCTGACCTCCGTC 615

QY 255 ACTGGGGGTGAATTCAGCTGCCGCTTCACTTCGGGAGGTCCAGAGCGTTGGTAG 314
Db 616 ACTGGGGGTGAATTCAGCTGCCGCTTCACTTCGGGAGGTCCAGAGCGTTGGTAG 675
QY 315 CCCGCTTACGATTCCTGTATCTCCAAAGTTGGGCTGTGACGCCATGAGGAGCTGCAAC 374
Db 676 CCCGCTTACGATTCCTGTATCTCCAAAGTTGGGCTGTGACGCCATGAGGAGCTGCAAC 735
QY 375 CAGAGGCTATTGACGCCATTCAGAGCAAGGCGCTGTTAGCAAGGCTGAGGAGCGCTGC 434
Db 736 CAGAGGCTATTGACGCCATTCAGAGCAAGGCGCTGTTAGCAAGGCTGAGGAGCGCTGC 795
QY 435 TGAGCAAGTGGATTCGACCAAGCCAGCCACTTTCAGAGACCTTCAGAGACCTGTGCA 494
Db 796 TGAGCAAGTGGATTCGACCAAGCCAGCCACTTTCAGAGACCTTCAGAGACCTGTGCA 855
QY 495 GACACCTGATGCTTCTTCACTGTGGCCGCTACCGGAGAGGCGCTGAGGCCACTGCGAC 554
Db 856 GACACCTGATGCTTCTTCACTGTGGCCGCTACCGGAGAGGCGCTGAGGCCACTGCGAC 915
QY 555 TCATGAAGCAATTAACCTGTGAGAGCAAGACAGTGCAGCCGCTGCGCAAGGGGAGC 614
Db 916 TCATGAAGCAATTAACCTGTGAGAGCAAGACAGTGCAGCCGCTGCGCAAGGGGAGC 975
QY 615 AAGTGTGAACCTTCTGTGATCAGAGCACTGATTTGATGACATGATGCTCAAGGACATGC 674
Db 976 AAGTGTGAACCTTCTGTGATCAGAGCACTGATTTGATGACATGATGCTCAAGGACATGC 1035
QY 675 GAGATGAGTCTTGGAAACATGAGCTGATGTGTGCTGACCGCGCCAGAGCGAGAGATTTC 734
Db 1036 GAGATGAGTCTTGGAAACATGAGCTGATGTGTGCTGACCGCGCCAGAGCGAGAGATTTC 1095
QY 735 GGCAGCTGGAACAGGAACCTGATTAAGTGGGAGTGAAGTGAAGACATCAAGGACATGA 794
Db 1096 GGCAGCTGGAACAGGAACCTGATTAAGTGGGAGTGAAGTGAAGACATCAAGGACATGA 1155
QY 795 GCTCTCCGGAATTGACCAACAGACACATGCGAGTGTGCGGCGCATGTGCGGTACC 854
Db 1156 GCTCTCCGGAATTGACCAACAGACACATGCGAGTGTGCGGCGCATGTGCGGTACC 1215
QY 855 TGATGCGCTGCGTGAATCACTCTGGGAGAGCAACCAAGATTAACGATTAATGTGG 914
Db 1216 TGATGCGCTGCGTGAATCACTCTGGGAGAGCAACCAAGATTAACGATTAATGTGG 1275
QY 915 ACCGTCTCTGGAAGGTCGCGGCTGGAAGATATCCCGAACAAGGTGTCAATCAAGCTGA 974
Db 1276 ACCGTCTCTGGAAGGTCGCGGCTGGAAGATATCCCGAACAAGGTGTCAATCAAGCTGA 1335
QY 975 AGAACCAAGGTGATTTCTTCAATGCAATGAGGTGCAAGGCCCATCTACATGATGAC 1034
Db 1336 AGAACCAAGGTGATTTCTTCAATGCAATGAGGTGCAAGGCCCATCTACATGATGAC 1395
QY 1035 GCGCGGTCTCTGTGCTCCAAATGCGCTCAGCAACCACTCTGTGTGAGATTCGCCA 1094
Db 1396 GCGCGGTCTCTGTGCTCCAAATGCGCTCAGCAACCACTCTGTGTGAGATTCGCCA 1455
QY 1095 GCTTCGCAATTTGCTCTTCTTATCAACCAAGACCTCATTTGCGCTCATCGGCTGAGGCTG 1154
Db 1456 GCTTCGCAATTTGCTCTTCTTATCAACCAAGACCTCATTTGCGCTCATCGGCTGAGGCTG 1515
QY 1155 CCAAGATCAGACCAAGTGAAGATGAGGAGGACTGTGAGGCGCTTCGCGGCGTGT 1214
Db 1516 CCAAGATCAGACCAAGTGAAGATGAGGAGGACTGTGAGGCGCTTCGCGGCGTGT 1575
QY 1215 CCCCTGCCACTCCAGCGCGCTTGAAGCTGGAATCAAGGCTCTGGAATAAACTGCGGAGT 1274
Db 1576 CCCCTGCCACTCCAGCGCGCTTGAAGCTGGAATCAAGGCTCTGGAATAAACTGCGGAGT 1635
QY 1275 GGGAGGCTCAGACTCGGGCCATTGATTTGAGCTTTGAGGAGGATGAGGCTGCGCTTTG 1334
Db 1636 GGGAGGCTCAGACTCGGGCCATTGATTTGAGCTTTGAGGAGGATGAGGCTGCGCTTTG 1695
QY 1335 TGAAGCCAGCAGAGGCTGAGAACTCAGAGGCTTCCTAGATCAAGAGCCCTCCCATCTT 1394

PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250391.
PR	01-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-541565/60.	
DR		
XX		
PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating nervous system	
PT	cancers and metastases -	
XX		
XX		
PS	Disclosure; SEQ ID NO 12345; 1701pp + Sequence Listing; English.	
XX		
CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins	
CC	(ABA1678-ABA18001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are	
CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful	
CC	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast	
CC	and ovarian cancer and other cancers of the adrenal gland, bone, bone	
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;	
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune	
CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	
CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative	
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias;	
CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and	
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal	
CC	and parasitic infections.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX		
SQ	Sequence 10579 BP; 2298 A; 2894 C; 3129 G; 2258 T; 0 other;	
Query Match	26.7%;	Score 414.4; DB 22; Length 10579;
Best Local Similarity	81.6%;	Pred. No. 1.1e-92;
Matches 560; Conservative	0; Mismatches	1; Indels 125; Gaps
958	AGGTCATCAAGCTGAAGAACACCGGTATTTCTTCATTGCGCAATGAGGTCACCGCC	10127
10005	AGGTCATCAAGCTGAAGAACACCGGTATTTCTTCATTGCGCAATGAGGTCACCGCC	9946

Oy		1018	CATTAATAAGTGAAGAACGGCCGCGTGCTGTCTGTTGGCTTCAAATTGGGCCCTCACAACAATC	1077		
Db		9945	CATTAATCATGAATGAACGGCCGGTAGTCTGTGTGCATCCAATGGCGCTCACCAACTC	9866		
Oy		1078	TGTGTGTG----- 	1085		
Db		9885	TGTGTGTGAAGTGTAGCTGGGGAGNAGACAGAAAGCCAGATGACCTTAAGCCGTGTGA	9826		
Oy		1086	----- 	1085		
Db		9825	GCCAGTACACCAAATTGTCTTGTCACCACCTGTCTTAAGCCAATTGTCTCCACC	9766		
Oy		1086	-AGATGCGCAGCGCTGGATTCGTCTTCATTACAACAGGACCTAATTGCCCTATCAGG	11444		
Db		9765	CAGATCCCACGCTCGCATTCGTCTTCTTAACAACAGAACCTATTGCCCTATCAGG	9706		
Oy		1145	GCTGAGGCTGCCAAGATCAACACCAAGTAGGAATGTGGCAGACCTGTGGGCCCTCTC	1204		
Db		9705	GCTGAGGCTGCCAAGATCAACACCAAGTAGGAATGTGGCAGACCTGTGGGCCCTCTC	9646		
Oy		1205	CGGCTGTTCCTCCCTGCACTCCAGCCCCCTTGAGCTGGGAACTGAGGCTCTTGAAAAA	1264		
Db		9645	CGGCTGTTCCTCCCTGCACTCCAGCCCCCTTGAGCTGGGAACTGAGGCTCTTGAAAAA	9586		
Oy		1265	CTTGAGGAGTGGGAGGCTGACGTGGGGGCAATTGAATTGAGCCTTGAAGGAGATAGG	1324		
Db		9585	CTTGAGGAGTGGGAGGCTGACGTGGGGGCAATTGAATTGAGCCTTGAAGGAGATAGG	9526		
Oy		1325	CTGAGCTTTGTGAAGCAGAGAGGCTGAGAACTCAGGCTTCCCTAGATCAGAGCCCC	1384		
Db		9525	CTGAGCTTTGTGAAGCAGAGAGGCTGAGAACTCAGGCTTCCCTAGATCAGAGCCCC	9466		
Oy		1385	TCGCCATCTTCTCTCTCTTAATAAACACCTACCCCCATTACCCCCCATTTGCCACT	1444		
Db		9465	TCGCCATCTTCTCTCTCTTAATAAACACCTACCCCCATTACCCCCCATTTGCCACT	9418		
Oy		1445	TCATCTCCTGTCACAGAGATTAGCCACAGACTTCTTTATTTGTTTTCTTTTSTA	1504		
Db		9417	TCATCTCCTGTCACAGAGATTAGCCCTCAGACTCTCTTTATTTGTTTTCTTTTSTA	9358		
Oy		1505	AATAAAAAGCAGCAGGTTCCAAAGTA	1530		
Db		9357	AATAAAAAGCAGCAGGTTCCAAAGTA	9332		
RESULT 4 AAK70045/c AAK70045 standard; DNA; 10579 BP.						
XX	ID	AAK70045				
XX	AC	AAK70045;				
XX	DT	06-NOV-2001 (first entry)				
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24857.					
XX	XX					
KW	Human; immune; haematopoietic; immunohaematopoietic antigen; cancer;					
KX	Cytosolic; gene therapy; vaccine; metastasis; ds.					
OS	Homo sapiens.					
PV	WO200157182-A2.					
PD	09-AUG-2001.					
PF	17-JAN-2001; 2001WO-US01354.					
PR	31-JAN-2000; 2000US-0179065.					
PR	04-FEB-2000; 2000US-0180628.					
PR	24-FEB-2000; 2000US-0184664.					
PR	02-MAR-2000; 2000US-0186350.					
PR	16-MAR-2000; 2000US-0189874.					
PR	17-MAR-2000; 2000US-0190076.					

CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AA64703
CC to AA87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AA54942 to AA54950 and AA82169
CC represent sequences used in the exemplification of the present invention.

XX
SQ Sequence 10579 BP; 2298 A; 2894 C; 3129 G; 2258 T; 0 other;

Query Match 26.7%; Score 414.4; DB 22; Length 10579;
Best Local Similarity 81.6%; Pred. No. 1.1e-92;
Matches 560; Conservative 0; Mismatches 1; Indels 125; Gaps 2;

Qy 958 AGGTGTCATCAAGCTGAAGAACACGCGTATTTCTTCATGTCATGAGGTCGACGGCC 1017
Db 10005 AGGTGTCATCAAGCTGAAGAACACGCGTATTTCTTCATGTCATGAGGTCGACGGCC 9946
Qy 1018 CATCAGCATCGATGAGCGCGCGTCTCTGTGGCTCCAAATGGCGCTCAGCAACATCTC 1077
Db 9945 CATCAGCATCGATGAGCGCGCGTCTCTGTGGCTCCAAATGGCGCTCAGCAACATCTC 9886
Qy 1078 TGTGTGTG----- 1085
Db 9885 TGTGTGTGAGGAGCTGGGGAGAGGAGGCAAGATGAGACCTAGGCGTGTGA 9826
Qy 1086 ----- 1085
Db 9825 GCCAATACCACTTGGTCTGTGTCACCTGTGTAAACCACTTGTCTCCACC 9766
Qy 1086 -AGATCGCAGCGCTGGAGTTGGTCTTCTTATCAACAGGACCTATGCGCTCATCAGG 1144
Db 9765 CAGATCGCAGCGCTGGAGTTGGTCTTCTTATCAACAGGACCTATGCGCTCATCAGG 9706
Qy 1145 GCTGAGCTGCGCAAGATCAACAAGTAGAGATGTGGCAGACTGATGGGCCCTCTC 1204
Db 9705 GCTGAGCTGCGCAAGATCAACAAGTAGAGATGTGGCAGACTGATGGGCCCTCTC 9646
Qy 1205 CGGCTGTTTCCCTGCGCATCTCAAGCCCCCTTGAAGTGGAACTCAAGCTCTTGAAAAA 1264
Db 9645 CGGCTGTTTCCCTGCGCATCTCAAGCCCCCTTGAAGTGGAACTCAAGCTCTTGAAAAA 9586
Qy 1265 CCGTGGCAGTGGAGGCTCAGCTGGGGCCATTGATTTGAGCCTTGGAGGAGATAGGG 1324
Db 9585 CCGTGGCAGTGGAGGCTCAGCTGGGGCCATTGATTTGAGGAGGAGATAGGG 9526
Qy 1325 CTGGCTTTGTGAAGCCAGCAGAGGCTGAGAACTCAGGCTTCCCTAGTCCAGAGCCC 1384
Db 9525 CTGGCTTTGTGAAGCCAGCAGAGGCTGAGAACTCAGGCTTCCCTAGTCCAGAGCCC 9466
Qy 1385 TCCCATCTTCTCTCTTAAAAACAACCTAACCCCCCATTTACCCCCCATTTGCACCT 1444
Db 9465 TCCCATCTTCTCTCTTAAAAACAAC-----CTACCCCCCATTTGCACCT 9418
Qy 1445 TCACCTCCGTGTCTCAGCTGATGAGCCCTGAGACTCTCTTTATATGTTTCTTTTGA 1504
Db 9417 TCACCTCCGTGTCTCAGCTGATGAGCCCTGAGACTCTCTTTATATGTTTCTTTTGA 9358
Qy 1505 AATAAAGACACAGATTCCAAGTA 1530
Db 9357 AATAAAGACACAGATTCCAAGTA 9332

RESULT 5
ABA20015/C
ID ABA20015 standard; DNA; 28995 BP.

XX
AC ABA20015;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 12346.
XX
KW Human; neurotropic; neuroprotective; cytosolic; dermatological; virocidic;
KW immunosuppressive; anti-inflammation; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisclerol; antianemic; antiarthritic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antileuk; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229347.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.

Db 19525 GCTGAGGCTGCCAAGATCAACACAGTAGAGAGTGTGGCAGAGCTCGTGGGCCCTCTC 19466
Qy 1205 CGGCGCTGTTTCCCTGACCACTCAGACCCCTTGTAGCTGGGAACCTCAGGCTCCTGGAAAA 1264
Db 19465 CGGCTGTTTCCCTGACCACTCAGACCCCTTGTAGCTGGGAACCTCAGGCTCCTGGAAAA 19406
Qy 1265 CCTGGGCAGTGGAGGCTCAGCTGGGGCCATTTGATTTGAGCCTTTGAGGAGAGATAGGG 1324
Db 19405 CCTGGGCAGTGGAGGCTCAGCTGGGGCCATTTGATTTGAGCCTTTGAGGAGAGATAGGG 19346
Qy 1325 CTGGCTTTTGTAGACCAAGAGGCTGAGAACTCAGGCTTCCCTAGATCCAGAGCCCC 1384
Db 19345 CTGGCTTTTGTAGACCAAGAGGCTGAGAACTCAGGCTTCCCTAGATCCAGAGCCCC 19286
Qy 1385 TCCCATCTTCTCTCTCTAAACACCCCTACCCCATTCACCCCATTTGACACT 1444
Db 19285 TCCCATCTTCTCTCTCTCTAAACACCC-----CTACCCCATTTGACACT 19238
Qy 1445 TCACCTCTGTGTCTCAGCTGATTAAGCTCAGACTCTTTTATTTGTTTCTTTGTA 1504
Db 19237 TCACCTCTGTGTCTCAGCTGATTAAGCTCAGACTCTTTTATTTGTTTCTTTGTA 19178
Qy 1505 AATAAAGACACAGGTTCCAAAGTA 1530
Db 19177 AATAAAGACACAGGTTCCAAAGTA 19152

RESULT 6
AAK70046/c
ID AAK70046 standard; DNA; 28995 BP.
XX
AC AAK70046;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24858.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cyostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225269.
PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232387.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236357.
PR 29-SEP-2000; 2000US-0236358.
PR 29-SEP-2000; 2000US-0236359.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM,
 XX
 DR WPI: 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 24858; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX Sequence 28995 BP; 6151 A; 8327 C; 8051 G; 6466 T; 0 other;
 SQ

Query Match 26.7%; Score 414.4; DB 22; Length 28995;
 Best Local Similarity 81.6%; Pred. No. 1.6e-92;

Matches 560; Conservative 0; Mismatches 1; Indels 125; Gaps 2;
 QY 958 AGGTGTCATCAGCTGAAGAACAAGTGATTTCTTCAATGCGAATGAGGGTGAAGCGCC 1017
 |||||
 Db 19825 AGGTGTCATCAGCTGAAGAACAAGTGATTTCTTCAATGCGAATGAGGGTGAAGCGCC 19766
 QY 1018 CATCTACATCGATGAGCGCCGCTCTGTGTGCTTCAATGCGCTCAGCAACAATC 1077
 |||||
 Db 19765 CATCTACATCGATGAGCGCCGCTCTGTGTGCTTCAATGCGCTCAGCAACAATC 19706
 QY 1078 TGTGTGTG----- 1085
 |||||
 Db 19705 TGTGTGTGAGTGAAGCTGGGAGAGAGAGCAAGGCCAGATGAGACTTGGTGTGA 19646
 QY 1086 ----- 1085
 Db 19645 GCCAGTACCAACTTGGTCTGTGTGCACTCTGTCTTAAAGCACTTGTCTCCACC 19586
 QY 1086 -AGATGCGCAGCTGCGATTCCTTCTTATCAACAGACCTGATTCCTCATCAGG 1144
 |||||
 Db 19585 CAGATGCGCAGCTGCGATTCCTTCTTATCAACAGACCTGATTCCTCATCAGG 19526
 QY 1145 GCTGAGCTGCCAAGATCAACCAAGTGAATGTGTGCGAGACTGTGGGCTCTTC 1204
 |||||
 Db 19525 GCTGAGCTGCCAAGATCAACCAAGTGAATGTGTGCGAGACTGTGGGCTCTTC 19466
 QY 1205 CGGCTGTTTCCCTGCGCACTGCGGCTTGTAGCTGAGGCTCCTGGAATA 1264
 |||||
 Db 19465 CGGCTGTTTCCCTGCGCACTGCGGCTTGTAGCTGAGGCTCCTGGAATA 19406
 QY 1265 CTTGGGCAAGTGGAGGCTGAGCTGCGGCTTGTAGCTTGTAGGAGGATAGGG 1324
 |||||
 Db 19405 CTTGGGCAAGTGGAGGCTGAGCTGCGGCTTGTAGCTTGTAGGAGGATAGGG 19346
 QY 1325 CTGGGCTTTGTAGAGCCAGAGAGGCTGAACCTGAGGCTTCTTGAATCCAGAGCCC 1384
 |||||
 Db 19345 CTGGGCTTTGTAGAGCCAGAGAGGCTGAACCTGAGGCTTCTTGAATCCAGAGCCC 19286
 QY 1385 TCCCATCTTCTCTCTCTTAAACAACCTTACCCCTTACCCCATTTGCAACT 1444
 |||||
 Db 19285 TCCCATCTTCTCTCTTAAACAACCTTACCCCTTACCCCATTTGCAACT 19238
 QY 1445 TCACCTCTGTGTCTCAGGTGATAGCTCAGACTTCTTATGTTTCTTTGTA 1504
 |||||
 Db 19237 TCACCTCTGTGTCTCAGGTGATAGCTCAGACTTCTTATGTTTCTTTGTA 19178
 QY 1505 AATTAAGACACAGGTTCCAAAGTA 1530
 |||||
 Db 19177 AATTAAGACACAGGTTCCAAAGTA 19152
 RESULT 7
 AAK79967
 ID AAK79967 standard; DNA; 28995 BP.
 XX
 AC AAK79967;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34779.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PF 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225256.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231966.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
DA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
FT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX

PS Disclosure; SEQ ID NO 34779; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM62169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 28995 BP; 6466 A; 8051 C; 8327 G; 6151 T; 0 other;

Query Match 26.7%; Score 414.4; DB 22; Length 28995;
Best Local Similarity 81.6%; Pred. No. 1,6e-92;
Matches 560; Conservative 0; Mismatches 1; Indels 125; Gaps 2;

QY 958 AGGTGTCATCAAGCTGAAAGAACACCGGTGATTTCTTCATTGCGCAATGAGGCTGAGAGGCC 1017
DB 9171 AGGTGTCATCAAGCTGAAAGAACACCGGTGATTTCTTCATTGCGCAATGAGGCTGAGAGGCC 9230
QY 1018 CATTCATCATGATGAGGAGCGCGGTCTGTGGCTCCAATGGCGGCTCAGCAACAATC 1077
DB 9231 CATTCATCATGATGAGGAGCGCGGTCTGTGGCTCCAATGGCGGCTCAGCAACAATC 9290
QY 1078 TGTGTGTG----- 1085
DB 9291 TGTGTGTGAGTGTAGCTGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9350
QY 1086 ----- 1085
DB 9351 GCCAGTACACCACTTGTCTGTGTCACCTTCTTAAGCCACTTTGTCTCCACC 9410
QY 1086 -AGATGAGGAGCGGCGGTGATTTCTTCATTGCGCAATGAGGCTGAGAGGCC 1144
DB 9411 CAGATGAGGAGCGGCGGTGATTTCTTCATTGCGCAATGAGGCTGAGAGGCC 9470
QY 1145 GCTGAGGCTGCAAGATCACCAACAGTGAAGATGTGGCAGAGCTCGTGGCCCTCTC 1204
DB 9471 GCTGAGGCTGCAAGATCACCAACAGTGAAGATGTGGCAGAGCTCGTGGCCCTCTC 9530
QY 1205 CGGCGTGTTCCTGCTGCACTCCAGCCCTTGAAGCTGGGAATCTAGGCTCTCTGAAAAA 1264
DB 9531 CGGCGTGTTCCTGCTGCACTCCAGCCCTTGAAGCTGGGAATCTAGGCTCTCTGAAAAA 9590
QY 1265 CTTGGGAGTGGAGGCTCAGCTGGCGGCATTGATTTGAGCTTTGAGGAGAGATAGGG 1324
DB 9591 CTTGGGAGTGGAGGCTCAGCTGGCGGCATTGATTTGAGCTTTGAGGAGAGATAGGG 9650
QY 1325 CTGGCTTTTGTAGGAGCAGAGAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCC 1384
DB 9651 CTGGCTTTTGTAGGAGCAGAGAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCC 9710
QY 1385 TCCCATCTTCTCTCTCTTAAAAACAACCTACCCCATTTCAACCCCATTTGCGCACT 1444
DB 9711 TCCCATCTTCTCTCTCTTAAAAACAACCTACCCCATTTCAACCCCATTTGCGCACT 9758
QY 1445 TCAGCTCTGTGCTCAGGCTGATTAAGCTCAGACTCTTTATTTATTTTCTTTTGA 1504
DB 9759 TCAGCTCTGTGCTCAGGCTGATTAAGCTCAGACTCTTTATTTATTTTCTTTTGA 9818
QY 1505 AATAAAGACACGAGGTTCCAAAGTA 1530
DB 9819 AATAAAGACACGAGGTTCCAAAGTA 9844

RESULT 8
ID AAK65213
AAK65213 standard; DNA; 28995 BP.
AC AAK65213;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40025.
DE Human immune/haematopoietic; immune/haematopoietic antigen; cancer;
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200157182-A2.
PD
XX 09-AUG-2001.
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209457.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225265.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227109.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231080.

Db 1746 CTACGAGCTGATCAACGCCATCCGCCAGAGAGCGCCCAAGA 1786
RESULT 10
ABL02114/c
ID ABL02114 standard; cDNA; 4027 BP.
XX ABL02114;
XX ABL02114;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 824.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX OS
XX MO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX PF
XX 23-MAR-2000; 2000US-191637P.
XX PR
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX DR
XX P-PSDB; ABB58011.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX
XX Claim 1; SEQ ID NO 824; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABR57737-ABR72072).
XX (ABR57737-ABR72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 4027 BP; 1010 A; 968 C; 1042 G; 1007 T; 0 other;
Query March 18.4%; Score 285.6; DB 23; Length 4027;
Best Local Similarity 57.6%; Pred. No. 8.3e-61;
Matches 611; Conservative 0; Mismatches 384; Indels 65; Gaps 3;
Qy 166 GGTGACCAAGATCTGGGCGGCTGGAAGCTGCAATGACCTCTGCTCATTAATGCTGT 225
Db 2179 GGGCCACCAAGATTTGGGCGGCTGGAAGCCCATTAACACTTTGGCCCTCATTAATGGCAT 2120
Qy 226 GTTGACCAACCAAGCTGACCTCCGCTCACTGGAGCGTGAATTAAGTGGCGCTTCAAC 285
Db 2119 CCAGCAGACCAACGATCTCGGATATCATCGCGCGTTAAGTTTCTCGCAAGTTTAC 2060
Qy 286 CCTTGAGAGGTCCAGGAGCGTTGGTAGCGCTGCTACATGATCTGTCATCTCAAGTT 345
Db 2059 TCTGAGAGAACTACAGCAGCGCTGGTAGCTCTGCTGATGAGCGCTGGGCTTCAAGAT 2000
Qy 346 GGCTGTGAGGCGCATGAGGCACTGACCCAGAGGCTATTGAGCGCATTCAGAGCAAGGC 405

Db 1999 TGCAGTGTCCGCATTAAGAAACCTGCACCCCGAACTCGTGGAGTCCCTTACGCCCAAGC 1940
Qy 406 CCGTTTAGCAAGGCTGAGAGCAGCTGCTGAGCAAGT----- 444
Db 1939 TTTGTACAGCGTCAGAGAGAGAGATCTGCTTGGCACTTAAGACCTTAAGTACTATT 1880
Qy 445 -----GGGATGACCAAGCCAGCCACC 466
Db 1879 TTCAATCATGATTTAAGCTTTTCTAATCATGATTAATTCATTCAGAGTGGAAACAAACAAA 1820
Qy 467 TTGAGACCTTCCAGAGACCTGCTGCAAGACACCTGATGCTTCTACCTGCGCCGCTACC 526
Db 1819 CTTGAGCAGTTCCAGAGCTACTGATTAAGATGCTCCGTTTCTACTGCGCCGCACT 1760
Qy 527 GCGAAGGCCCTGACGGCCCACTGGAGAGCTATGAAAGAGTAATTCTGCTGAGAGACAG 586
Db 1759 GCGAATCTTGGCAAAACATTTGGCTTTCTCAAGAGTACACCTCTGTTCCGGATAG 1700
Qy 587 ACAGTGCAGCCGCT---GCCCAAGGGGACCAGTGTGAACTTCTGTATGACAGAGAC 643
Db 1699 TCAGTGAAGCCTATATATATGAGCAGGATCAGAGCACTCAGCTTCTGGAGTGAAGAT 1640
Qy 644 CTGATTGATGACAGTACCTCAAGACATGCGAGATGAGTCTTGGAAACATGAGCTGATG 703
Db 1639 CAAATCTTGCAGACGACCTGAAAGCAGTACGAAAGCTCTGAGATGAGGAGACT 1580
Qy 704 GTGCTGACCGGCGCCCAAGACGAGATTTGGCAGCTGGAACAGAACTGCAATAGTG 763
Db 1579 CTGGAGATCGTGCMAAAGATATATTTGCTGCTGGAGAAAGAGACTTCCGCGCTGG 1520
Qy 764 CAGGTGCTAGTGAACAGCATCAGAGCATGAGCTCTCCG---GACTTGCACCAACAGACA 820
Db 1519 GCGGTCTGTGAGATTCGCTGCTTAACTGACGCGCTGCTCGAGTTGACACACAGACA 1460
Qy 821 CTGGCAGTGTGCGGGGCGCATGAGTGCCTGATGCTGCTGCGCTGAGATCACCTG 880
Db 1459 CTGGCTGCTGTGTGGCGCGCATGTGCGTAACTGATGCTTCCMAAGAAATTAATATT 1400
Qy 881 GGCAGAGCAACCAAGGATTAACAGATGATGAGACCTGTCTGAGAGGTCCGGCGCTGG 940
Db 1399 GGTGCGATGCGCAAGAGACTGTGTGAGATGATGATGAGGCTGAGAGACCGGCTGCG 1340
Qy 941 AAGATATCCCGAAACCAAGGTGCTATCAAGCTGAAAGAAACAGGATTTCTTCAATGCC 1000
Db 1339 AAGATCTCTCCCGCAAGAAACATTAAGCTGCGCAGCAATGGGATTTCTTCAATGCT 1280
Qy 1001 AATGAGGTGACGCGCCATCTAATCATGAGAGCGCGGTCTGTGCTTCCAAATGG 1060
Db 1279 AACGAGGAAAGAGGCGCATCTTATGACGCGCACTCCTTGTATCCGCCAACAAAGCT 1220
Qy 1061 CGCTTCAGCAACACTCTGTGTGAGATTCGCCAGCTGCGATTTGTTCTTATCAAC 1120
Db 1219 CGACTCGGTCACAACTGACAGGTGAATATCGGTGCGCTTCACTTCTGTGTCAAC 1160
Qy 1121 CAGGACCTCATTTGCCCTGAGGCTGAGGCTGCAAGA 1160
Db 1159 TACGAGCTGATCAACGCGCATCCGCCAGAGAGCGCCAGA 1120
RESULT 11
AAK79966
ID AAK79966 standard; DNA; 8204 BP.
XX AAK79966;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34778.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytoblastic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.

XX
PM WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226881.
PR 22-AUG-2000; 2000US-0226881.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.

PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239393.
PR 13-OCT-2000; 2000US-0239393.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241825.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0255678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 34778; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 8204 BP; 1865 A; 2157 C; 2229 G; 1953 T; 0 other;
Query Match 10.3%; Score 159.4; DB 22; Length 8204;
Best Local Similarity 99.4%; Pred. No. 3e-29; Mismatches 0; Gaps 0;
Matches 160; Conservative 0; Indels 1;
QY 72 AGGATCCAAAGCCCCGACACTCTGTGCCACCCAGCCCGCCCTGACTCA 131
Db 3245 AGGATCCAAAGCCCCGACACTCTGTGCCACCCAGCCCGCCCTGACTCA 3304
QY 132 CCAAGCGGTGAAGAAGTAACAGCCACTTCAGGTGACCAAGATCTGGGCGCTGGA 191
Db 3305 CCAAGCGGTGAAGAAGTAACAGCCACTTCAGGTGACCAAGATCTGGGCGCTGGA 3364
QY 192 AGCCTGCAATGACCTCTGCTATTAATGCTGTGTCAG 232
Db 3365 AGCCTGCAATGACCTCTGCTATTAATGCTGTGTCAG 3405
RESULT 12
AAK85212
ID AAK85212 standard; DNA; 8204 BP.
XX
AC AAK85212;
XX
DT 07-NOV-2001 (first entry)
XX
DB Human immune/hematopoietic antigen genomic sequence SEQ ID NO:40024.
XX
KW Human; immune; hematopoietic; immune/hematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
FN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218230.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227109.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249399.
PR 17-NOV-2000; 2000US-0249399.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Disclosure; SEQ ID NO 40024; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic

CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 8204 BP, 1865 A; 2157 C; 2229 G; 1953 T; 0 other;

Query Match 10.3%; Score 159.4; DB 22; Length 8204;
Best Local Similarity 99.4%; Pred. No. 3e-29; Mismatches 1; Indels 0; Gaps 0;
Matches 160; Conservative 0;

QY 72 AGGATCCAAAGCCCCCAGCACTCTGTGCCACCCAGCCCAAGCCCTTGACTCA 131
DB 3245 AGGTATCCAAAGCCCCCAGCACTCTGTGCCACCCAGCCCAAGCCCTTGACTCA 3304
QY 132 CCNAGCGTGTGAAGAAGTAAACAGCCACTTCAGGTGACCAAGATCTGGGCGCTGGA 191
DB 3305 CCNAGCGTGTGAAGAAGTAAACAGCCACTTCAGGTGACCAAGATCTGGGCGCTGGA 3364
QY 192 AGCCTGCAATGACCTCTGCTCATTAATGCTGTGTTGCA 232
DB 3365 AGCCTGCAATGACCTCTGCTCATTAATGCTGTGTTGCA 3405

RESULT 13
AAAT23953
ID AAAT23953 standard; cDNA to mRNA; 144 BP.
XX
XX AAAT23953;
AC
XX
XX 27-AUG-1996 (first entry)
DT
XX
XX Human gene signature HUMG505899.
DE
XX
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
XX Homo sapiens.
OS
XX
XX W09514772-A1.
PN
XX
XX 01-JUN-1995.
PD
XX
XX I1-NOV-1994; 94WO-JP01916.
PF
XX
XX 12-NOV-1993; 93JP-0355504.
PR
XX
XX (MATS/) MATSUBARA K.
PA (OKUBO/) OKUBO K.
XX
XX
PI Matsubara K, Okubo K;
XX
XX WPI; 1995-206931/27.
XX
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
XX
XX Claim 1; Page 1495; 2245bp; Japanese.

CC	A single-stranded DNA (or its complementary strand or the corresp.
CC	double-stranded DNA) which comprises one of the 7637 "GS" sequences
CC	given in AAT19001-T26837 and which is able to hybridise to part of
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC	sequences were obtained from 3'-directed cDNA libraries prepared
CC	from various human tissues; synthesis of cDNA was initiated from the
CC	3'-end of mRNA by using poly(17) as the sole primer. Since the 3'-
CC	untranslated sequence is unique to a particular mRNA species, almost
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC	is constructed so as to reflect accurately the relative abundance of
CC	different mRNAs in the particular tissue from which it was derived.
CC	The appearance frequency of a given GS in a cDNA library can be
CC	determined (esp. using primers and probes derived from the GS
CC	sequences) as a means of diagnosing abnormal cell function or for
CC	recognising different cell types.
SQ	Sequence 144 BP; 34 A; 50 C; 15 G; 44 T; 1 other;
Query Match	7.8%; Score 121; DB 16; Length 144;
Best Local Similarity	91.7%; Pred. No. 2e-20;
Matches 143; Conservative	0; Mismatches 1; Indels 12; Gaps 1;
OY	13372 GATTCAGACAGCCCCCTCCCACATCTTCCCTGTAAATAACAACCCTAACCCCACTTACCC 1431
DB	1 GATTCAGACAGCCCCCTCCCACATCTTCCCTGTAAATAACAAC-----CTACCC 48
OY	14332 CCCATTGCACACTTCACCTCCTGTCCTCCAGCTGATTAGCTCAGACTCTTTATTG 1491
DB	49 CCCATTGCCACCTTCACCTCTGTGTCTCGAGCTGATTAGCTCAGACTCTTTATTG 108
OY	1492 TTTTCTTTTGTAAATAAANAAGCACCGATTCCA 1527
DB	109 TTTTCTTTTGTAAATAAANAAGCACAGTCCA 144
RESULT 14	
ID	AAK87783
XX	AAK87783 standard, cDNA; 457 BP.
AC	AAK87783;
DT	05-NOV-2001 (first entry)
DE	Human digestive system antigen coding sequence SEQ ID NO: 99.
XX	Human, digestive system antigen; gene therapy; cancer; appendicitis;
KM	ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW	digestive system disorder; Meckel's diverticulum; ss.
XX	Homo sapiens.
OS	WO200155314-A2.
PN	02-AUG-2001.
PD	17-JAN-2001; 2001WO-USO1324.
XX	
PF	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0215135.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.

[illegible]

XX	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246509.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249221.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Rosen CA, Barash SC, Ruben SM;	
P1	WPI: 2001-502630/55.	
DR	P-PSDB; AAM92010.	
XX		
PT	Polynucleotides encoding digestive system antigens, useful for	
PT	diagnosing, treating, preventing and/or prognosing disorders of the	
PT	digestive system, particularly cancer and cancer metastases -	
PS		
XX	Claim 1; SEQ ID NO 99; 986bp; English.	
XX		
CC	The present invention provides the protein and coding sequences of a	
CC	number of human digestive system antigens. These can be used in the	
CC	diagnosis, treatment and prevention of digestive system disorders,	
CC	including cancer, Meckel's diverticulum, bacterial or parasitic	
CC	infections, appendicitis, Hirschsprung's disease, chronic colitis or	
CC	ulcerative colitis. The present sequence is a cDNA encoding a digestive	
CC	system antigen of the invention.	
XX		
SQ	Sequence 457 BP; 94 A; 129 C; 123 G; 109 T; 2 other;	
Query Match	4.6%; Score 71.6; DB 22; Length 457;	
Best Local Similarity	94.4%; Pred. No. 6; 5e-08;	

Matches	85;	Conservative	0;	Mismatches	4;	Indels	1;	Gaps	1
Oy	958	AGGTCGTCATCAACGTCGAGACCAACGCGATTCTTCATTCGTCATGAGGGGTCGACGGCC	1017						
Db	276	AGGTTCATCAACGTCGAGAACCAACGGTGATTTCTTCATTCGCAATGAGGGTCGACGC	334						
Oy	1018	CATTCATCATCGATGAGCGCGCGTGCTCTG	1047						
Db	335	CATTCATCATGATGAGCGCGCGTGCTCTG	364						
RESULT 15									
AA331700									
AA331700	standard; cDNA; 457 BP.								
AA331700;									
DT									
XX	04-DEC-2001 (first entry)								
DE	Human liver associated cDNA polynucleotide #20.								
XX									
KW	Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;								
KW	chicken; sheep; immunosuppressive; antirheumatic; vasotropic;								
KW	antirheumatic; antiproliferative; cytoskeletal; cardiant; neuroprotective;								
KW	cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;								
KW	ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;								
KW	hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;								
KW	cerebrovascular disorder; nervous system disorder; bacterial infection;								
KW	fungal infection; viral infection; ocular disorder; endocrine disorder;								
KW	gastrointestinal disorder; renal disorder; respiratory disorder;								
KW	wound healing; skin aging; organ transplantation; tissue regeneration;								
KW	anti-fertility.								
XX									
OS	Homo sapiens.								
XX									
PN	MO20015355-A1.								
XX									
PD	02-AUG-2001.								
XX									
PF	17-JAN-2001; 2001MO-US01351.								
XX									
PR	31-JAN-2000; 2000US-0179065.								
PR	04-FEB-2000; 2000US-0180628.								
PR	24-FEB-2000; 2000US-0184664.								
PR	02-MAR-2000; 2000US-0186350.								
PR	16-MAR-2000; 2000US-0189874.								
PR	17-MAR-2000; 2000US-0190076.								
PR	18-APR-2000; 2000US-0198123.								
PR	19-MAY-2000; 2000US-0205515.								

Db 276 AGGTGTCATCAAGCTGAGAGACACGGTGATTTCTTCATTGCCAATGAGGGTGGACGG-C 334

Oy 1018 CATCTACATCGATGAGACGGCCGGTGTCTTG 1047

Db 335 CATCTACATCGATGAGACGGCCGGTGTCTTG 364

Search completed: March 27, 2003, 05:40:40
Job time : 632 secs

• • •

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 05:29:20 ; Search time 81 Seconds
(without alignments)
5879.864 Million cell updates/sec

Title: US-09-744-125A-1
Perfect score: 1553
Sequence: 1 gaattcgccagcagtgctggg.....aaaaaaaaaacctcgag 1553

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.4	3.2	1926	US-09-249-585A-2	Sequence 2, Appli
2	49.4	3.2	2580	US-09-050-863-2	Sequence 2, Appli
3	49.4	3.2	2580	US-09-359-081-2	Sequence 2, Appli
4	49.4	3.2	5452	US-09-130-114-1	Sequence 1, Appli
5	49.4	3.2	9600	US-08-910-647-1	Sequence 1, Appli
6	49.4	3.2	9600	US-09-620-925-1	Sequence 1, Appli
7	49.4	3.2	10596	US-07-884-811-15	Sequence 15, Appli
8	49.4	3.2	10596	US-07-885-971-15	Sequence 15, Appli
9	49.4	3.2	10596	US-08-087-783A-15	Sequence 15, Appli
10	49.4	3.2	10596	US-08-194-088B-15	Sequence 15, Appli
11	49.4	3.2	10596	US-08-194-087-15	Sequence 15, Appli
12	49.4	3.2	10596	PCT-US93-04648-15	Sequence 15, Appli
13	48.6	3.1	3489	US-08-728-323A-1	Sequence 1, Appli
14	48.6	3.1	3489	US-09-298-568-1	Sequence 1, Appli
15	48.6	3.1	33207	US-08-770-379-20	Sequence 20, Appli
16	48.6	3.1	33207	US-08-757-669A-20	Sequence 20, Appli
17	48.6	3.1	33207	US-09-230-371A-20	Sequence 20, Appli
18	47.6	3.1	16442	US-08-781-891-208	Sequence 208, App
19	46	3.0	7218	US-08-232-463-14	Sequence 14, Appli
20	45.6	2.9	913	US-08-217-327-3	Sequence 3, Appli
21	45.6	2.9	913	US-07-885-970A-3	Sequence 3, Appli
22	45.6	2.9	913	US-08-298-687A-3	Sequence 3, Appli
23	45.6	2.9	913	US-08-530-797-2	Sequence 2, Appli
24	45.6	2.9	913	US-08-298-829-3	Sequence 3, Appli
25	45.6	2.9	913	US-08-787-335-2	Sequence 2, Appli
26	45.6	2.9	1984	US-07-885-970A-25	Sequence 25, Appli
27	45.6	2.9	1985	US-08-298-687A-25	Sequence 25, Appli

C	28	45.6	2.9	1985	1	US-08-298-829-25	Sequence 25, Appli
C	29	45.6	2.9	7218	1	US-08-232-463-14	Sequence 14, Appli
C	30	42.6	2.7	2338	1	US-08-425-069-1	Sequence 1, Appli
C	31	42.6	2.7	2338	2	US-08-317-844B-1	Sequence 1, Appli
C	32	42.6	2.7	2793	1	US-08-209-747-1	Sequence 1, Appli
C	33	42	2.7	2793	1	US-08-458-298-1	Sequence 1, Appli
C	34	41.4	2.7	397	3	US-09-253-691-3	Sequence 3, Appli
C	35	41	2.6	432	1	US-08-642-255-48	Sequence 48, Appli
C	36	41	2.6	756	1	US-08-642-255-50	Sequence 50, Appli
C	37	39.8	2.6	51259	3	US-08-781-891-209	Sequence 209, App
C	38	39.6	2.5	1508	4	US-09-039-046-1	Sequence 1, Appli
C	39	39.6	2.5	5267	3	US-08-976-255-2	Sequence 2, Appli
C	40	38.8	2.5	2830	2	US-09-010-928B-1	Sequence 1, Appli
C	41	38.8	2.5	9551	1	US-08-056-200-93	Sequence 93, Appli
C	42	38.8	2.5	9551	2	US-08-800-644-93	Sequence 93, Appli
C	43	38.4	2.5	9757	1	US-08-093-453B-1	Sequence 1, Appli
C	44	38.4	2.5	9759	1	US-08-459-041A-1	Sequence 1, Appli
C	45	38.4	2.5	9759	3	US-08-999-733-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-249-585A-2
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match 3.2%; Score 49.4; DB 4; Length 1926;
Best Local Similarity 49.8%; Pred. No. 0.0015;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY	527	GGGAAGGCGCTGCGAGGCGCCACATGCGAGCTCATGAAGCAGTAATTACTGTGAGGACGAG	586
DB	359	GGGCGAGGCGGCGAGGCGAGGAGGAGGCGGCGAGGCGAGGAGGCGGCGGCGGCGG	418
QY	587	ACAGTCAGCGCGCTGCGCCAAAGGAGGAGCCAGTGTGAATCTTCTGTATGACAGGACCTG	646
DB	419	GAGGCGGCGAGGCGAGGAGGAGGAGGCGGCGAGGAGGCGGCGGCGGCGGCGGCGG	478
QY	647	ATTGATGACAGTAACTTAAGGACATGCGAGATGAGTCTTGTGAACATGAGCTGATGTG	706
DB	479	GAGGCGGCGGCGAGGCGGCGAGGAGGCGGCGAGGCGGCGAGGAGGCGGCGGCGGCGG	538
QY	707	GCTGCGCGGCGCGGAGGAGGAGGAGGAGTTCGCGAGCTGGAACGAACTGCATTAAGTGGC	766
DB	539	GGGCGAGGCGGCGGCGAGGAGGAGGAGGAGGCGGCGAGGCGGCGGCGGCGGCGGCGG	598
QY	767	GTGCTAGTGGGA	777
DB	599	GAGGCGGCGAGGA	609

RESULT 2
US-09-050-863-2
; Sequence 2, Application US/09050863

```
; Patent No. 611411
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; US-09-050-863-2

Query Match          3.2%; Score 49.4; DB 3; Length 2580;
Best Local Similarity 49.8%; Pred. No. 0.0017;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCCCTGCAAGCCCACTGCGACGCTCATGACGATATTACTGCTGAGAGCACCG 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 742 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 587 ACAGTGCACCGCTGCCCAAGAGGAGCAAGTGTGAACCTTCTGTGATGACGACCTG 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 802 GAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGACATGAGCTGATGTG 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 862 GAGGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 707 GCTGACCGGCGCCAGAAAGCAGAGATTGCGCAGCTGGAACAGAACTGATAGTGGCAG 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 922 GGGCAGAGAGGGGCGAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 767 GTGCTAGTGA 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 982 GAGGGGCGAGGA 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 3
US-09-359-081-2
; Sequence 2, Application US/09359081
```

```
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
```

```
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/359,081
; FILING DATE: 22-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/050,863
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-09-359-081-2

Query Match          3.2%; Score 49.4; DB 4; Length 2580;
Best Local Similarity 49.8%; Pred. No. 0.0017;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCCCTGCAAGCCCACTGCGACGCTCATGACGATATTACTGCTGAGAGCACCG 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 742 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 587 ACAGTGCACCGCTGCCCAAGAGGAGCAAGTGTGAACCTTCTGTGATGACGACCTG 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 802 GAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGACATGAGCTGATGTG 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 862 GAGGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 707 GCTGACCGGCGCCAGAAAGCAGAGATTGCGCAGCTGGAACAGAACTGATAGTGGCAG 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 922 GGGCAGAGAGGGGCGAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 767 GTGCTAGTGA 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 982 GAGGGGCGAGGA 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 4
US-09-130-114-1/c
; Sequence 1, Application US/09130114
```

```
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dama, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
```

TITLE OF INVENTION: From Multiple Transfected Epismes
FILE REFERENCE: 0867/ID903US1
CURRENT APPLICATION NUMBER: US/09/130.114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 1
LENGTH: 5452
TYPE: DNA
ORGANISM: VEBNA
US-09-130-114-1

Query Match 3.2%; Score 49.4; DB 2; Length 5452;
Best Local Similarity 49.8%; Pred. No. 0.0025;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGCCCTGACAGCCCACTGACGATCATGAACGATATTACTGTGAGAGACCG 586
DB 2063 GGGCGAGAGGGGCGAGACGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAG 2004
QY 587 ACAGTGCAGCCGCTGCCCAAGAGGAGCCAAAGTGTGAATTCTTGATGACAGAGCTG 646
DB 2003 GAGGGGCGAGAGCAGAGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAG 1944
QY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTTGGAACATGAGCTGATG 706
DB 1943 GAGGAGGGGCGAGAGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAG 1884
QY 707 GCTGACCGGCGCCAGAGCGAGAGATTGGCAGCTGGAACAGGAAGTGCATTAAGTG 766
DB 1883 GGGCGAGAGGGGCGAGAGAGAGAGGGGCGAGAGAGGGGCGAGAGAGAGAGAG 1824
QY 767 GTGCTAGTGA 777
DB 1823 GAGGGGCGAGAG 1813

RESULT 5

US-08-910-647-1
Sequence 1, Application US/08910647
Patent No. 6251433
GENERAL INFORMATION:
APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for
TITLE OF INVENTION: Polynucleotide Delivery
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910.647
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

Query Match 3.2%; Score 49.4; DB 4; Length 9600;
Best Local Similarity 49.8%; Pred. No. 0.0033;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGCCCTGACAGCCCACTGAGACTCATGAACGATATTACTGTGAGAGACCG 586
DB 788 GGGCGAGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAG 847
QY 587 ACAGTGCAGCCGCTGCCCAAGAGGAGCCAAAGTGTGAATTCTTGATGACAGAGCTG 646
DB 848 GAGGGGCGAGAGCAGAGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAG 907
QY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTTGGAACATGAGCTGATG 706
DB 908 GAGGAGGGGCGAGAGGGGCGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAG 967
QY 707 GCTGACCGGCGCCAGAGCGAGAGATTGGCAGCTGGAACAGGAAGTGCATTAAGTG 766
DB 968 GGGCGAGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAG 1027
QY 767 GTGCTAGTGA 777
DB 1028 GAGGGGCGAGAG 1018

RESULT 6

US-09-620-925-1
Sequence 1, Application US/09620925
Patent No. 6468986
GENERAL INFORMATION:
APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for
TITLE OF INVENTION: Polynucleotide Delivery
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/620.925
FILING DATE: 21-Jul-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/910.647
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-620-925-1

Query Match 3.2%; Score 49.4; DB 4; Length 9600;
Best Local Similarity 49.8%; Pred. No. 0.0033;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGCCCTGACAGCCCACTGCGACGTCATGAAGCAGTATTACTGCTGGAGACCG 586
DB 788 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
QY 587 ACAGTCAGCCCTGCGCCAAAGGAGCAAGTCGTAACCTCTGTATGACAGAGACCTG 646
DB 848 GAGGGGCGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 907
QY 647 ATTGATGACAGTAACTCAAGGACATGCGAGTGAAGTCTTGGAAACATGAGCTGATG 706
DB 908 GAGGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 967
QY 707 GCTGACCCGCGCAGAGGAGGAGATTGCGGACGCTGGAACAGGAACCTGATAGTGCGAG 766
DB 968 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027
QY 767 GTGCTAGTGA 777
DB 1028 GAGGGGCGACGA 1038

RESULT 7

US-07-884-811-15
; Sequence 15, Application US/07884811
; Patent No. 5316921

; GENERAL INFORMATION:

APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

; COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

NAME: Dregger, Ginger R.
REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: 755.1

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881

; TELEK: 910/371-7168

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10596 bases

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

US-07-884-811-15

Query Match 3.2%; Score 49.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 0.0035;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGCCCTGACAGCCCACTGCGACGTCATGAAGCAGTATTACTGCTGGAGACCG 586
DB 2322 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2381
QY 587 ACAGTCAGCCCTGCGCCAAAGGAGCAAGTCGTAACCTCTGTATGACAGAGACCTG 646
DB 2382 GAGGGGCGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2441
QY 647 ATTGATGACAGTAACTCAAGGACATGCGAGTGAAGTCTTGGAAACATGAGCTGATG 706
DB 2442 GAGGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2501
QY 707 GCTGACCCGCGCAGAGGAGGAGATTGCGGACGCTGGAACAGGAACCTGATAGTGCGAG 766
DB 2502 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2561
QY 767 GTGCTAGTGA 777
DB 2562 GAGGGGCGACGA 2572

RESULT 8

US-07-885-971-15
; Sequence 15, Application US/07885971
; Patent No. 5328837

; GENERAL INFORMATION:

APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

; COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/885,971
FILING DATE: 19920518

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

NAME: Dregger, Ginger R.
REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: 779

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881

; TELEK: 910/371-7168

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10596 bases

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

US-07-885-971-15

Query Match 3.2%; Score 49.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 0.0035;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGCCCTGACAGCCCACTGCGACGTCATGAAGCAGTATTACTGCTGGAGACCG 586
DB 2322 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2381

OY	707	GCTCACCGGGCCCGAAGCCAGAGAATTGCGCACTGGAACAGGAATCGATAAGTGGCAG	766
Dδ	2502	GCGCAGGAGGGGCAGGAGCAGAGAGAGGGGCAGAGAGGGGCAGGAGGCGAGGAG	2562
OY	767	GTCTACTGGA	777
Dδ	2562	GAGGGGCGAGG	2572

US-08-194-087-15
US-08-194-087-15
Sequence 15. Application US/08194087
Patent No 5879910
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: HEPHYOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,087
FILING DATE: 18-MAY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 779
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1056 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-194-087-15

[illegible]

Qy	767	GTGCTAGTGA	777
Db	2562	GAGGGCAGGA	2572

RESULT 12
 PCT-US93-04648-15
 Sequence 15, Application PC/TUS9304648
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc., Goddard, Paul J, Lokker, Natalie A., Mark, Melanie R
 TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: patin (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/04648
 FILING DATE: 19930517
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/884811
 FILING DATE: 18-MAY-92
 PRIOR APPLICATION DATA: 07/885971
 APPLICATION NUMBER: 07/885971
 FILING DATE: 18-MAY-92
 ATTORNEY/AGENT INFORMATION:
 NAME: Dreger, Ginger R.
 REGISTRATION NUMBER: 33,055
 REFERENCE/DOCKET NUMBER: 755,779P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-3216
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10596 bases
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 PCT-US93-04648-15

Query Match	Similarity	3.2%	Score 49.4	DB 5	Length 10596
Best Local	Similarity 49.8%	Pred. No. 0.0035			
Matches 125	Conservative 0	Mismatches 126	Indels 0	Gaps 0	
QY	527	GGCAAGGCCCTTCAGAGCCCACTGCGAGCTCAATGAACAGATATTACTGCTGGAGACCAAG	586		
DB	2322	GGGCGAGAGGGGCGAGAGCAGAGAGAGAGGGGCGAGGACAGAGAGAGGGGCGAGAGGGGCGAG	2381		
QY	587	ACAGTCAGACCGCTGCGCCAAAGGGGACCAAGTGTGAACCTTCTTGATGCAGAGGACCTG	646		
DB	2382	GAGGGGCGAGGACAGAGAGAGGGGCGAGGACGAGGAGGAGGGGCGAGGGGCGAGGAGCAG	2441		
QY	647	ATTGATGACAGTAACTCAAGAGCATGCGAGATGAAGGTCCTTGGAAACATGAGCTGATGGTG	706		
DB	2442	GAGGAGGGGCGAGAGAGGGGCGAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGAGG	2501		
QY	707	GCTGACCGGCGCCGAGAAACCGAGAGATTGGCGAGCTGTGSAACAGGAACGTGATTAAGTGGCAG	766		
DB	2502	GGGCGAGAGGGGCGAGAGCAGAGGAGGAGGGGCGAGGAGGGGCGAGAGGCGCAGAGGACGAGG	2561		
QY	767	GTGCTAGTGGGA	777		
DB	2562	GAGGGGCGAGGA	2572		

RESULT 13

US-08-728-323A-1

Sequence 1, Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

TITLE OF INVENTION: Immediate Early Protein From Kaposi's

TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/728,323A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3489 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..3489

US-08-728-323A-1

Query Match 3.1%; Score 48.6; DB 2; Length 3489;

Best Local Similarity 49.8%; Pred. No. 0.0033;

Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 531 AGCCCTGAGGCCCACTGCGAGCTCATGAAGCAGATTACTCTGAGGACCGACAG 590
DB 1901 AGCAGCAGCAGAGATGACGACGACGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 1960
QY 591 TGCAGCCCGTCCCAAGGGGACCAAGTGTGAATCTTCTGTGATGACAGAGCCTGATTG 650
DB 1961 AGCAGCAGCAGAGATGACGACGACGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 2020
QY 651 ATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGAACATGAGCTGATGTGGCTG 710
DB 2021 AGCAGCAGCAGAGATGAGCAGCAGCAGATGAGCAGCAGAGATGAGCAGCAGAGATG 2080
QY 711 ACCGGCGCCAGAGAGAGAGATTTCGCGAGCTGGAACAGGAACGATAGTGGCAGGTGC 770
DB 2081 AGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 2140
QY 771 TAGTGA 777
|||

DB 2141 AGCAGGA 2147

RESULT 14

US-09-298-568-1

Sequence 1, Application US/09298568

Patent No. 6322792

GENERAL INFORMATION:

APPLICANT: Kieff, Elliott D.

APPLICANT: Ballester, Mary E.

APPLICANT: Kaye, Kenneth M.

TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO

TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE

FILE REFERENCE: 16412-1001R

CURRENT APPLICATION NUMBER: US/09/298,568

EARLIER FILING DATE: 1998-04-21

EARLIER APPLICATION NUMBER: US 60/109,422

NUMBER OF SEQ ID NOS: 3

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 1

LENGTH: 3489

TYPE: DNA

ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-1

Query Match 3.1%; Score 48.6; DB 4; Length 3489;

Best Local Similarity 49.8%; Pred. No. 0.0033;

Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 531 AGCCCTGAGGCCCACTGCGAGCTCATGAAGCAGATTACTCTGAGGACCGACAG 590
DB 1901 AGCAGCAGCAGAGATGACGACGACGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 1960
QY 591 TGCAGCCCGTCCCAAGGGGACCAAGTGTGAATCTTCTGTGATGACAGAGCCTGATTG 650
DB 1961 AGCAGCAGCAGAGATGACGACGACGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 2020
QY 651 ATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGAACATGAGCTGATGTGGCTG 710
DB 2021 AGCAGCAGCAGAGATGACGACGACGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 2080
QY 711 ACCGGCGCCAGAGAGAGAGATTTCGCGAGCTGGAACAGGAACGATAGTGGCAGGTGC 770
DB 2081 AGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 2140
QY 771 TAGTGA 777
|||
DB 2141 AGCAGGA 2147

RESULT 15

US-08-770-379-20/c

Sequence 20, Application US/08770379

Patent No. 5849564

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED

TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-770-379-20

Query Match      3.1%; Score 48.6; DB 2; Length 32207;
Best Local Similarity 49.8%; Pred. No. 0.0099;
Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 531 AGGCCCTGAGGCCCACTGGCAGCTCATGAAAGCAGTATTACCTGCTGAGAGACCAGACAG 590
DB 20096 AGCAGCAGCAGAGTAGAGCAGCAGCAGCAGAGTAGACGACGACGAGTAGACAGCAGAGTAG 20037
QY 591 TGCAGCCGCTGCCCAAGGGGACCAAGTGTGAATTCTTCTGATGACAGAGACCTGATTG 650
DB 20036 AGCAGCAGCAGAGTAGAGCAGCAGCAGCAGAGTAGACGACGACGAGTAGACAGCAGAGTAG 19977
QY 651 ATGACAGTAAGCTCAAGGACATGCGAGATGAGGTCCTGGAAATGAGCTGATGATGGCTG 710
DB 19976 AGCAGCAGCAGAGTAGAGCAGCAGCAGCAGAGTAGACGACGACGAGTAGACAGCAGAGTAG 19917
QY 711 ACCGGGCGCCAGAAGCGAGATTGGCAGCTGGAAACAGAACTGCAATAAGTGGCAGGTGC 770
DB 19916 AGCAGCAGCAGAGTAGAGCAGAGCAGCAGAGTAGACGACGAGAGCAGCAGATGAGCAGAGAC 19857
QY 771 TAGTGGA 777
DB 19856 AGCAGGA 19850
```

Search completed: March 27, 2003, 07:50:43
Job time : 223 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 07:46:17 ; Search time 186 Seconds
(without alignments)
7105.167 Million cell updates/sec

Title: US-09-744-125A-1
Perfect score: 1553
Sequence: 1 gaatccgcagcagcaggtcgg99.....aaaaaaaaaacctcag 1553

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214.6	13.8	371	10	US-09-783-590-2176 Sequence 2176, Ap
2	157.8	10.2	318	10	US-09-783-590-10575 Sequence 10575, A
3	101.2	6.5	385	10	US-09-783-590-10146 Sequence 10146, A
4	71.6	4.6	457	10	US-09-764-887-30 Sequence 30, Appl
5	58	3.7	2108	10	US-09-962-832-225 Sequence 225, App
6	46.6	3.0	8459	10	US-09-817-913-8 Sequence 8, Appl1
7	46.6	3.0	8459	10	US-09-817-538-8 Sequence 8, Appl1
8	45.4	2.9	3131	10	US-09-817-913-14 Sequence 14, Appl
9	45.4	2.9	3131	10	US-09-817-538-14 Sequence 14, Appl
10	45.2	2.9	14800	10	US-09-954-456-1601 Sequence 1601, Ap
11	43.4	2.8	7305	9	US-10-156-239-9 Sequence 9, Appl1
12	43.4	2.8	7305	9	US-09-795-693-9 Sequence 9, Appl1
13	43.4	2.8	8056	9	US-10-072-621-3 Sequence 3, Appl1
14	43.4	2.8	8195	9	US-10-156-239-7 Sequence 7, Appl1
15	43.4	2.8	8195	10	US-09-785-693-7 Sequence 7, Appl1
16	42.4	2.7	2553	10	US-09-815-242-7690 Sequence 7690, Ap
17	42.2	2.7	422	10	US-09-854-133-337 Sequence 337, App
18	42.2	2.7	422	10	US-09-738-973-337 Sequence 337, App
19	42.2	2.7	765	9	US-09-925-299-143 Sequence 143, App

20	42	2.7	765	10	US-09-925-299-143 Sequence 143, App
21	42	2.7	1852	10	US-09-969-852-4 Sequence 24, Appl1
22	40.6	2.6	485	9	US-09-992-598-245 Sequence 245, App
23	40.6	2.6	485	9	US-09-989-293A-245 Sequence 245, App
24	40.6	2.6	485	9	US-10-063-547-43 Sequence 43, Appl
25	40.6	2.6	485	9	US-09-989-735-245 Sequence 245, App
26	40.6	2.6	485	9	US-09-990-444-245 Sequence 245, App
27	40.6	2.6	485	9	US-09-989-730-245 Sequence 245, App
28	40.6	2.6	485	9	US-09-990-436-245 Sequence 245, App
29	40.6	2.6	485	9	US-09-991-181-245 Sequence 245, App
30	40.6	2.6	485	9	US-09-993-687-245 Sequence 245, App
31	40.6	2.6	485	9	US-09-989-734-245 Sequence 245, App
32	40.6	2.6	485	9	US-09-997-653-245 Sequence 245, App
33	40.6	2.6	485	9	US-10-174-590-185 Sequence 185, App
34	40.6	2.6	485	9	US-10-176-758-185 Sequence 185, App
35	40.6	2.6	485	9	US-10-063-616-43 Sequence 43, Appl
36	40.6	2.6	485	9	US-10-175-737-185 Sequence 185, App
37	40.6	2.6	485	9	US-09-993-667-245 Sequence 245, App
38	40.6	2.6	485	9	US-10-063-502-43 Sequence 43, Appl
39	40.6	2.6	485	9	US-10-173-706-185 Sequence 185, App
40	40.6	2.6	485	9	US-10-175-738-185 Sequence 185, App
41	40.6	2.6	485	9	US-10-175-752-185 Sequence 185, App
42	40.6	2.6	485	9	US-10-176-482-185 Sequence 185, App
43	40.6	2.6	485	9	US-10-176-757-185 Sequence 185, App
44	40.6	2.6	485	9	US-10-176-913-185 Sequence 185, App
45	40.6	2.6	485	9	US-10-180-552-185 Sequence 185, App

ALIGNMENTS

RESULT 1
US-09-783-590-2176
Sequence 2176, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillion, Patrick J.
APPLICANT: Haeseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2176
LENGTH: 371
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (106)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (114)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (117)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (125)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (131)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (146)


```
RESULT 3
US-09-783-590-10146
; Sequence 10146, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haeseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10146
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (237)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (309)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (340)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (352)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (379)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (385)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-10146

Query Match          6.5%; Score 101.2; DB 10; Length 385;
Best Local Similarity 95.1%; Pred. No. 7.9e-21;
Matches 136; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
```

```
; Sequence 30, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA113
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 30
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (387)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-887-30
```

```
Query Match          4.6%; Score 71.6; DB 10; Length 457;
Best Local Similarity 94.4%; Pred. No. 1e-11;
Matches 85; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Oy 958 AGGTGTCATCAAGCTGAAGACAACAGGATTTCTTCATGCCATGAGGTCGACGGCC 1017
Db 276 AGGTGTCATCAAGCTGAAGACAACAGGATTTCTTCATGCCATGAGGTCGACGG-C 334

Oy 1018 CATCTACATCATGATGACGGCGGCTGCTCTG 1047
Db 335 CATCTACATCATGATGACGGCGGCTGCTCTG 364
```

```
RESULT 5
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225
```

```
Query Match          3.7%; Score 58; DB 10; Length 2108;
Best Local Similarity 53.0%; Pred. No. 3.3e-07;
Matches 124; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Oy 536 CTGCAGGCGCCACTGGCACTGATGAAGACATTAACCTGCTGAGAGACAGACAGTGCG 595
Db 1056 CTGAAGCACTGAGACAGACAGAGGGGAGCTGGAGCACTGGAGACACAGAGAGGGCAG 1115

Oy 596 CCGTGCCCAAGAGGAGCAAGAGTGTGAATCTCTGTATGACAGAGACCTGATTGATGAC 655
Db 1116 CTGGGGCTCCAGAGCAGCAGAGTGTGCTGAGAGCAGCTTAAAGAAACACAGAGGGCAG 1175

Oy 656 AGTAAGCTCAAGAGACATGCGATGAGGTCTTGGAACATGAGCTGATGTGCTGACCGG 715
Db 1176 CCAAGCACTGAGAGAGAGAGGGGAGCTGAAGCACTGTGTGCACAGAGAGGGGAG 1235
```

QY 716 CGCCAGAGCGAGATTGCGAGCTGGAACAGAACTGCATTAAGTGGCAGG 769
DB 1236 CTGAAGCATGTGTGACGACGAGGGGCGACTGTGACGACGAGAGGAGGAGT 1289

RESULT 6

US-09-817-913-8
; Sequence 8, Application US/09817913
; Patent No. US2002061860A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
; FILE REFERENCE: 106101.145
; CURRENT APPLICATION NUMBER: US/09/817,913
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 8459
; TYPE: DNA
; ORGANISM: Human
US-09-817-913-8

Query Match 3.0%; Score 46.6; DB 10; Length 8459;
Best Local Similarity 49.8%; Pred. No. 0.0022;
Matches 118; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 532 GGGCTGCGAGCGCCACTGCGAGCTCATGAAGCATTTACTGCTGAGAGACCAAGCACT 591
DB 1020 GGGCTGCGAGCGCCACTGCGAGCTCATGAAGCATTTACTGCTGAGAGACCAAGCACT 1079
QY 592 GCAGCGCGCTGCCCAAGGGGACCAAGTGTGAACTTCTGTGATGACAGAGACCTGATGA 651
DB 1080 GCAGCGCGCTGCCCAAGGGGACCAAGTGTGAACTTCTGTGATGACAGAGACCTGATGA 1139
QY 652 TGACAGTAAGCTCAAGAGCATGTGAGATGAGTCTGGAACATGAGCTGATGAGTGA 711
DB 1140 GGAGATGCTGGCATGACACACAGCAGAGGCTGTGGAACACACAGCGAAGCTGGAGAG 1199
QY 712 CCGGGGCCGAGAGCGAGATTGCGAGCTGGAACAGAACTGCATTAAGTGGCAGGT 768
DB 1200 GCACCGCGAGAGCGAGATTGCGAGCTGGAACAGAACTGCATTAAGTGGCAGGT 1256

RESULT 7

US-09-817-538-8
; Sequence 8, Application US/09817538
; Patent No. US20020137162A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
; FILE REFERENCE: 106101.144
; CURRENT APPLICATION NUMBER: US/09/817,538
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 8459
; TYPE: DNA
; ORGANISM: Human
US-09-817-538-8

Query Match 3.0%; Score 46.6; DB 10; Length 8459;
Best Local Similarity 49.8%; Pred. No. 0.0022;

Matches 118; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 532 GGGCTGCGAGCGCCACTGCGAGCTCATGAAGCATTTACTGCTGAGAGACCAAGCACT 591
DB 1020 GGGCTGCGAGCGCCACTGCGAGCTCATGAAGCATTTACTGCTGAGAGACCAAGCACT 1079
QY 592 GCAGCGCGCTGCCCAAGGGGACCAAGTGTGAACTTCTGTGATGACAGAGACCTGATGA 651
DB 1080 GCAGCGCGCTGCCCAAGGGGACCAAGTGTGAACTTCTGTGATGACAGAGACCTGATGA 1139
QY 652 TGACAGTAAGCTCAAGAGCATGTGAGATGAGTCTGGAACATGAGCTGATGAGTGA 711
DB 1140 GGAGATGCTGGCATGACACACAGCAGAGGCTGTGGAACACACAGCGAAGCTGGAGAG 1199
QY 712 CCGGGGCCGAGAGCGAGATTGCGAGCTGGAACAGAACTGCATTAAGTGGCAGGT 768
DB 1200 GCACCGCGAGAGCGAGATTGCGAGCTGGAACAGAACTGCATTAAGTGGCAGGT 1256

RESULT 8

US-09-817-913-14
; Sequence 14, Application US/09817913
; Patent No. US2002061860A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
; FILE REFERENCE: 106101.145
; CURRENT APPLICATION NUMBER: US/09/817,913
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 3131
; TYPE: DNA
; ORGANISM: Human
US-09-817-913-14

Query Match 2.9%; Score 45.4; DB 10; Length 3131;
Best Local Similarity 51.8%; Pred. No. 0.003;
Matches 103; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 712 CCGGGGCCGAGAGCGAGATTGCGAGCTGGAACAGAACTGCATTAAGTGGCAGGT 771
DB 1796 CCGAGGCGGAGAGGCTCCCTGGAAGAGCTGCACTGAGCGGCGAGTGT 1855
QY 772 AGTGAACAGATCAAGGACATGAGTCTCCGACTTCGACACACAGACCTGGCAGTGT 831
DB 1856 CCTTACGCGACACCAACCGGCTAGCGGCTCAACTGGAACAGGAGAGTGGCAGGCT 1915
QY 832 GCGGGGCCGATGTGCGGTAAGTGTGAGTGTGCGGTGAGATCACTGGGCGAGCAAC 891
DB 1916 CCGGACAGCGGAGATGTTGAGATGCTGCTGTGTGGGTGGGTGGAGACAGTGAAC 1975
QY 892 CAAGATTAACGAGATTGAT 910
DB 1976 CATCTGAATGAGCTTCAT 1994

RESULT 9

US-09-817-538-14
; Sequence 14, Application US/09817538
; Patent No. US20020137162A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
; FILE REFERENCE: 106101.144

;; CURRENT APPLICATION NUMBER: US/09/817,538
;; CURRENT FILING DATE: 2001-03-26
;; PRIOR APPLICATION NUMBER: US 60/192,157
;; PRIOR FILING DATE: 2000-03-24
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 14
;; LENGTH: 3131
;; TYPE: DNA
;; ORGANISM: Human
US-09-817-538-14

Query Match 2.9%; Score 45.4; DB 10; Length 3131;
Best Local Similarity 51.8%; Pred. No. 0.003;
Matches 103; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 712 CCGGCGCCAGAGCGAGATTCGGCAGCTGGAACAGAACTGCTAAGTGGCAGTGTCT 771
DB 1796 CCGAGGCGCGAAGGCTCTCCCTGGAAGAGCTGCACTGCTCCACTGTGAGCGGCACTGTCT 1855
QY 772 AGTGACAGACATCACAGCATGAGCTCTCCGAGCTTGACACACAGACACTGGAGTGTCT 831
DB 1856 CCTTACGAGCACCACCCGCTCAGCCGCTCAAACTGACAAAGGAGCTGGGAGGCT 1915
QY 832 GCGGCGCCGATGCTGCGGATCTGATGCGCTCGCGTGAATCAACCTGGGCGACAGCAAC 891
DB 1916 CTTGGCAGCGGATGTTTGAGATCTGCTGCTGTGGGTGGGATGACACTGACAC 1975
QY 892 CAAGATTAACCAATTGAT 910
DB 1976 CATCTGGAATGAGCTTCAT 1994

RESULT 10
US-09-954-456-1601
; Sequence 1601, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:

; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1601
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1601

Query Match 2.9%; Score 45.2; DB 10; Length 14800;

Best Local Similarity 46.1%; Pred. No. 0.008;
Matches 189; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

QY 360 TGAGGAGCTGCAACCCAGAGGCTATTTCAGCCATCCAGAGCAAGGCCCTGTTAGCAAG 419
DB 7535 TGCTGAGCTCAAGTCTGAGAGATGAGACGGTGCAGAGACAGCTCTCGAGAGGA 7594
QY 420 CTGAGAGCAGCTGCTGAGCAAAAGT---GGAGTCAGCCAGCCAGCCACTTGGAGACT 476
DB 7595 CGAGGCCCTGCGACAAAGCTTCTCTGAAAAGACAGCTCTACAGGGAGCGCT 7654
QY 477 TCCAGACTTGTGACAGACACCTGATGCTTCTACCTGGCCGTAACGCGAAGGCC 536
DB 7655 TCATGAGCAGAGAAAGCCAGAGCTGAGCAGCTCTTCCAGAGAGAGTGGCCAAAGGCAC 7714
QY 537 TGCAGGCCCACTGCGACCTCATGAGAGATTAACCTGCTGAGAGACCAAGCACTGACAG 596
DB 7715 AGCAGCTGCTGAGAGACAGACGGGAGCAGCAGAGATGAGACAGGAACGGAGCGGC 7774
QY 597 CGCTGCCCAAGGGGACCAAGTGTGAATCTTCTGATGCGAGACCTGATTGATGACA 656
DB 7775 TGCTGCTCAGCATGAGAGAGCGCGCGCGGCGAGCATGAGGCCGAGAGGCGCTGCGGC 7834
QY 657 GTAAGCTCAAGGACATGCGAGATGAGTCTTGAACATGAGCTGATGCTGACCGGC 716
DB 7835 GCAAGCAGAGAGAGCTGAGAGCTGAGAGCAGACCGCGGACAGAGAGAGCTGCTGG 7894
QY 717 GCCAGAACGAGAGATTCGGCAGCTGGAACAGAACTGCAATAGTGGCAG 766
DB 7895 CTGAGAGAACCAAGGCTGCTGAGCAGCTGACGCTCTGAGAGACAG 7944

RESULT 11
US-10-156-239-9
; Sequence 9, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:

; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Transl
; TITLE OF INVENTION: Atpase Molecule, A Human Ubiquitin Hydrolyase-Like Molecule, A Hur
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 7305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-156-239-9

Query Match 2.9%; Score 43.4; DB 9; Length 7305;
Best Local Similarity 45.3%; Pred. No. 0.019;
Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACAAACACAGACACTGCGAGTGTGCGGGCCGCGATGGTGGTACCTGATGCGC 862
Db 2902 GCCATGAGAGAGCGCGCTTTAGAGAGACCCGTGGCATGAGAGAGAGCCCAACCTG 2961
QY 863 TCGCGTAGATCACCTTGGGAGAGCAACCAAGATTAACAGATTGATGAGACTGTCT 922
Db 2962 CCTGTGTTGTGTGGTGAGCAAACTCAACCAAGTCTCAAGAGACGACAAAGAGCTGCGC 3021
QY 923 CTGAGAGGTCGGCCTGGAAGATATCCCGGAAAACAAGTGTATCAAGCTGAAGAACAC 982
Db 3022 CTGAACAAGCTAGCTTAACCTTAAGAAACAGAGTGTCTCTTCTTGGGCAACAC 3081
QY 983 GGTGATTTCTTATGTCATGAGAGGTGACCGCCATATCATGATGAGAGCGCGGTG 1042
Db 3082 GGGGGGGGAAACACCAACCACTGTCATCTGACCGGCTGTCCCTCAACGTGCGGT 3141
QY 1043 CTGTGTGCTTCCAAATGCGGCTCAGCAACAACTGTGTGAGATGCGCAGCTGCGA 1102
Db 3142 TCCGCCACCATCTAGGGGACAGACATCCGACGAGATGATGATGATGATGATGATG 3201
QY 1103 TTGCTCTTCTTATCAACCAAGACCTCATTTGCCCTTCATGAGGCTGAGG 1151
Db 3202 GGCATGTGCGCGACAGACATGTGCTTTTGAACGCGCTCACGCTGAGG 3250

RESULT 12

US-09-795-693-9
; Sequence 9, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; TITLE OF INVENTION: 32613, No. US20020068710A1 Human Transporters
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 7305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-693-9

Query Match 2.8%; Score 43.4; DB 10; Length 7305;

Best Local Similarity 45.3%; Pred. No. 0.019;

Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACAAACACAGACACTGCGAGTGTGCGGGCCGCGATGGTGGTACCTGATGCGC 862
Db 2902 GCCATGAGAGAGCGCGCTTTAGAGAGACCCGTGGCATGAGAGAGAGCCCAACCTG 2961
QY 863 TCGCGTAGATCACCTTGGGAGAGCAACCAAGATTAACAGATTGATGAGACTGTCT 922
Db 2962 CCTGTGTTGTGTGGTGAGCAAACTCAACCAAGTCTCAAGAGACGACAAAGAGCTGCGC 3021
QY 923 CTGAGAGGTCGGCCTGGAAGATATCCCGGAAAACAAGTGTATCAAGCTGAAGAACAC 982
Db 3022 CTGAACAAGCTAGCTTAACCTTAAGAAACAGAGTGTCTCTTCTTGGGCAACAC 3081
QY 983 GGTGATTTCTTATGTCATGAGAGGTGACCGCCATATCATGATGAGAGCGCGGTG 1042
Db 3082 GGGGGGGGAAACACCAACCACTGTCATCTGACCGGCTGTCCCTCAACGTGCGGT 3141
QY 1043 CTGTGTGCTTCCAAATGCGGCTCAGCAACAACTGTGTGAGATGCGCAGCTGCGA 1102
Db 3142 TCCGCCACCATCTAGGGGACAGACATCCGACGAGATGATGATGATGATGATGATG 3201
QY 1103 TTGCTCTTCTTATCAACCAAGACCTCATTTGCCCTTCATGAGGCTGAGG 1151
Db 3202 GGCATGTGCGCGACAGACATGTGCTTTTGAACGCGCTCACGCTGAGG 3250

RESULT 13

US-10-072-621-3
; Sequence 3, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-3

Query Match 2.8%; Score 43.4; DB 9; Length 8056;

Best Local Similarity 45.3%; Pred. No. 0.02;

Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACAAACACAGACACTGCGAGTGTGCGGGCCGCGATGGTGGTACCTGATGCGC 862
Db 2952 GCCATGAGAGAGCGCGCTTTAGAGAGACCCGTGGCATGAGAGAGAGCCCAACCTG 3011
QY 863 TCGCGTAGATCACCTTGGGAGAGCAACCAAGATTAACAGATTGATGAGACTGTCT 922
Db 3012 CCTGTGTTGTGTGGTGAGCAAACTCAACCAAGTCTCAAGAGAGAGAGAGAGTGGCGC 3071
QY 923 CTGAGAGGTCGGCCTGGAAGATATCCCGGAAAACAAGTGTATCAAGCTGAAGAACAC 982
Db 3072 CTGAACAAGCTAGCTTAACCTTAAGAAACAGAGTGTCTCTTCTTGGGCAACAC 3131
QY 983 GGTGATTTCTTATGTCATGAGAGGTGACCGCCATATCATGATGAGAGCGCGGTG 1042
Db 3132 GGGGGGGGAAACACCAACCACTGTCATCTGACCGGCTGTCTCTTCAACGTGCGGT 3191
QY 1043 CTGTGTGCTTCCAAATGCGGCTCAGCAACAACTGTGTGAGATGCGCAGCTGCGA 1102
Db 3192 TCCGCCACCATCTAGGGGACAGACATCCGACGAGATGATGATGATGATGATGATG 3251
QY 1103 TTGCTCTTCTTATCAACCAAGACCTCATTTGCCCTTCATGAGGCTGAGG 1151
Db 3252 GGCATGTGCGCGACAGACATGTGCTTTTGAACGCGCTCACGCTGAGG 3300

RESULT 14

US-10-156-239-7
; Sequence 7, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. US20030036074A1 Nucleic Acid Sequences Encoding Human Trans
; TITLE OF INVENTION: ATPase Molecule, A Human Ubiquitin Hydrolyase-Like Molecule, A Hu
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24

;; PRIOR APPLICATION NUMBER: 09/808,568
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 60/191,790
;; PRIOR FILING DATE: 2000-03-24
;; PRIOR APPLICATION NUMBER: 09/808,767
;; PRIOR FILING DATE: 2001-03-15
;; PRIOR APPLICATION NUMBER: 60/191,781
;; PRIOR FILING DATE: 2000-03-24
;; NUMBER OF SEQ ID NOS: 60
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7
;; LENGTH: 8195
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (132)...(7442)
;; US-10-156-239-7

Query Match 2.8%; Score 43.4; DB 9; Length 8195;
Best Local Similarity 45.3%; Pred. No. 0.021; Mismatches 191; Indels 0; Gaps 0;
Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Qy 803 GACTTCGACAAACGACACGCTGAGTCTGGGGCCGCGATGTCGGTACCTGATGCGC 862
Db 3033 GCCATGAGAGCCCGCGCTTTGAGAGAGCCGTGGCATGAGAGAGCCACCCACTG 3092
Qy 863 TCGGCTGATCACCCTGGGCGACGACCAAGATAACGATTGATGAGACTGTCT 922
Db 3093 CCTCTGTTGCTGTGTGAGACAACTCACAAGGTCTCAAGAGACGACAAAGCTGGCC 3152
Qy 923 CTGAGAGGTCCGGCCTGGAAGATATCCCGAAACAAAGTGTCTCAAGCTGAAGAACAC 982
Db 3153 CTGAACAAAGCTGAGCTGAACCTTAACGAGAACAGGTGTCTCTTGGGCCAACAC 3212
Qy 983 GGTGATTTCTTCAATTCGAATGAGGTCGAGGCCCATCTACATGATGAGCGCGGTG 1042
Db 3213 GGGGCGGCGAAGACCAACCATGTCATCTGACCGGCTGTCCCTCAACGTGGGT 3272
Qy 1043 CTGTGTGCTCCAAATGCGGCTCAGCAACACTGTGTGTGAGATGCGACGCTGCGA 1102
Db 3273 TCCGCAACATCTACGGGCAAGACATCCGACGAGATGATGAGATCCGCAAGAACTG 3332
Qy 1103 TTGCTTCTTATCAACGAGACCTCATGCGCTCATGAGGCTGAGG 1151
Db 3333 GGCATGTCCCGCAGACCAATGTGCTTTGACCGGCTCAGCGTGAAG 3381

RESULT 15
US-09-795-693-7
Sequence 7, Application US/09795693
Patent No. US20020068710A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
FILE REFERENCE: 35800/209292
CURRENT APPLICATION NUMBER: US/09/795,693
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/185,906
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 8195
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (132)...(7442)
US-09-795-693-7

Query Match 2.8%; Score 43.4; DB 10; Length 8195;

Best Local Similarity 45.3%; Pred. No. 0.021;
Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Qy 803 GACTTCGACAAACGACACCTGCGAGTGTGGGGCCGCGATGTCGGTACCTGATGCGC 862
Db 3033 GCCATGAGAGCCCGCGCTTTGAGAGAGCCGTGGCATGAGAGAGCCACCCACTG 3092
Qy 863 TCGGCTGATCACCCTGGGCGAGCAACCAAGATAACGATTGATGAGACTGTCT 922
Db 3093 CCTCTGTTGCTGTGTGAGACAACTCACAAGGTCTCAAGAGACGACAAAGCTGGCC 3152
Qy 923 CTGAGAGGTCCGGCCTGGAAGATATCCCGAAACAAAGTGTCTCAAGCTGAAGAACAC 982
Db 3153 CTGAACAAAGCTGAGCTGAACCTTAACGAGAACAGGTGTCTCTTGGGCCAACAC 3212
Qy 983 GGTGATTTCTTCAATTCGAATGAGGTCGAGGCCCATCTACATGATGAGCGCGGTG 1042
Db 3213 GGGGCGGCGAAGACCAACCATGTCATCTGACCGGCTGTTCCTCCCAACGTGGGT 3272
Qy 1043 CTGTGTGCTCCAAATGCGGCTCAGCAACACTGTGTGTGAGATGCGACGCTGCGA 1102
Db 3273 TCCGCAACATCTACGGGCAAGACATCCGACGAGATGATGAGATCCGCAAGAACTG 3332
Qy 1103 TTGCTTCTTATCAACGAGACCTCATGCGCTCATGAGGCTGAGG 1151
Db 3333 GGCATGTCCCGCAGACCAATGTGCTTTGACCGGCTCAGCGTGAAG 3381

Search completed: March 27, 2003, 09:16:22
Job time : 272 secs

This Page Blank (uspto)

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 180 a 275 c 273 g 227 t 11 others
ORIGIN

Query Match 59.5%; Score 924.2; DB 9; Length 966;
Best Local Similarity 97.8%; Pred. No. 1.9e-162;
Matches 945; Conservative 10; Mismatches 9; Indels 2; Gaps 2;

QY 448 ATGACACGACGACCCACCTTGGAGACTTCCAGACCTGCTGCACAGACACCCCTGATGC 507
DB 966 ATGACACGACGACCCACCTTGGAGACTTCCAGACCTGCTGCACAGACACCCCTGATGC 507
QY 508 CTCTTACCTGCGCCGTAACCGGAGCCCTGCAAGGCGCCCACTGCACTGATGAAGCACTA 567
DB 906 CTCTTACCTGCGCCGTAACCGGAGCCCTGCAAGGCGCCCACTGCACTGATGAAGCACTA 567
QY 568 TTACCTGCTGAGGAGCAGACAGTGCAGCGCGCCCAAGGGGACCAAGTGTGAACCTT 627
DB 846 TTACCTGCTGAGGAGCAGACAGTGCAGCGCGCCCAAGGGGACCAAGTGTGAACCTT 627
QY 628 CTCTGATGACAGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 687
DB 786 CTCTGATGACAGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 727
QY 688 GGAACATGAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 747
DB 726 GGAACATGAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 747
QY 748 GGAACATGAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 807
DB 666 GGAACATGAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 807
QY 808 CGACCAAC-CAGACACTGAGAGTGTGCGGGCGGACGATGCTGATGCTGATGCTGATGCT 866
DB 606 CGACCAACCAAGACACTGAGAGTGTGCGGGCGGACGATGCTGATGCTGATGCTGATGCT 866
QY 867 GTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
DB 546 GTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
QY 926 GAGGGGCGGCGCTGAGAGATATCCCGGAAACAGAGTGTATCAAGCTGAAAGAACAGG 985
DB 486 GAGGGGCGGCGCTGAGAGATATCCCGGAAACAGAGTGTATCAAGCTGAAAGAACAGG 985
QY 986 GATTTCTTATGTCATGAGGCTGACGCGCCATCTACATGATGAGACGCGCGGCTCTC 1045
DB 426 GATTTCTTATGTCATGAGGCTGACGCGCCATCTACATGATGAGACGCGCGGCTCTC 1045
QY 1046 TGTGGCTCCAATGCGGCTCAGCAACAATCTGTGTGTGAGATGCCAGGCTTGCATTC 1105
DB 366 TGTGGCTCCAATGCGGCTCAGCAACAATCTGTGTGTGAGATGCCAGGCTTGCATTC 1105
QY 1106 GTCTTCTTATCAACAGGACCTCATTTGCTCATAGGGCTGAGGCTGCAAGATCAAC 1165
DB 306 GTCTTCTTATCAACAGGACCTCATTTGCTCATAGGGCTGAGGCTGCAAGATCAAC 1165
QY 1166 CCACAGTGAAGATGAGTGGAGACTGTGTGGGCTCTCCGCGCTGTTCCTCCCTGCACT 1225
DB 246 CCACAGTGAAGATGAGTGGAGACTGTGTGGGCTCTCCGCGCTGTTCCTCCCTGCACT 1225
QY 1226 CCAGCCCCCTTGAAGCTGGGAATCTGAGGCTCTGGAAGAACCTGGGAGATGGAGGCTCAG 1285
DB 186 CCAGCCCCCTTGAAGCTGGGAATCTGAGGCTCTGGAAGAACCTGGGAGATGGAGGCTCAG 1285
QY 1286 CTGCGGGGCAATGATTTGAGGCTTTGAGGAGATGGGGTGGCTTTTGGAGCCAGCA 1345

DB 126 CTGCGGGGCAATGATTTGAGGCTTTGAGGAGATAGGAGCTGCGCTTTGTGAAGCAGCA 67
QY 1346 GAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCTCCCACTTCTCTCTCTA 1405
DB 66 GAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCTCCCACTTCTCTCTCTA 1405
QY 1406 AATACAA 1411
DB 6 AAAAAA 1

RESULT 2
AL578980/c 914 bp mRNA linear EST 16-FEB-2001
LOCUS AL578980 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DK012YC23 3
DEFINITION prime, mRNA sequence.
ACCESSION AL578980
VERSION AL578980.1 GI:12943576
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 914)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source location/Qualifiers
1..914
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DK012YC23"
/clone_id="LTI_NFL006.PL2"
/tissue_type="Placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 177 a 271 c 261 g 204 t 1 others
ORIGIN

Query Match 57.8%; Score 898.4; DB 9; Length 914;
Best Local Similarity 99.6%; Pred. No. 1.2e-157;
Matches 910; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 490 GACACAGACCTTATGCTTCTTACCTGCGCTGCGGAGGCGCCGACAGCCCACTG 549
DB 914 GACACAGACCTTATGCTTCTTACCTGCGCTGCGGAGGCGCCGACAGCCCACTG 549
QY 550 GCAGCTCATGAAGAGTATTAACCTGCTGAGAGACAGACAGTGCAGCCGCTGCCAAGG 609
DB 854 GCAGCTCATGAAGAGTATTAACCTGCTGAGAGACAGACAGTGCAGCCGCTGCCAAGG 609
QY 610 GGA-CGAGTGTGAACCTTCTGTATGACAGAGACCTGATTTATGACAGTTAAGCTCAAG 668
DB 794 GGA-CGAGTGTGAACCTTCTGTATGACAGAGACCTGATTTATGACAGTTAAGCTCAAG 668
QY 669 ACATGGAGATGAGTCTCTGGAACATGAGCTGATGAGTGTGCTGACCGGCGCGAAGCGAG 728
DB 734 ACATGGAGATGAGTCTCTGGAACATGAGCTGATGAGTGTGCTGACCGGCGCGAAGCGAG 728
QY 729 AGATTGCGAGCTGGAACAGAACTGATTAAGTGCAGGCTGCTAGTGAACAGATCAAG 788

```

Db 674 AGATTCCGACCTGGAGACAGAACTGATTAAGTGGACAGTGTCTAGTGAACAGCATCACAG 615
Qy 789 GCATGAGCTCTCCGAACTTTCAGCAACAGACACTGGAGTGGCGGGCCGATGTGC 848
Db 614 GCATGAGCTCTCCGAACTTTCAGCAACAGACACTGGAGTGGCGGGCCGATGTGC 555
Qy 849 GGTAAGCTGATGCGCTCGCGTGAAGATCACTGGGAGCAACCAAGATTAACCAAGATTG 908
Db 554 GGTAAGCTGATGCGCTCGCGTGAAGATCACTGGGAGCAACCAAGATTAACCAAGATTG 495
Qy 909 ATGTGAGACTGTCTTGTGAAGGCTCGGCGCTGGAAGATATCCCGGAAACAAGTGTATCA 968
Db 494 ATGTGAGACTGTCTTGTGAAGGCTCGGCGCTGGAAGATATCCCGGAAACAAGTGTATCA 435
Qy 969 AGCTGAAGAACAAAGGATTTCTTCAATTCGCAATGAGAGGTGACAGCGCCATCTACATCG 1028
Db 434 AGCTGAAGAACAAAGGATTTCTTCAATTCGCAATGAGAGGTGACAGCGCCATCTACATCG 375
Qy 1029 ATGAGACGGCCGGTGTCTGTGGCTTCAAAATGGCGCTCAGCAACAACTGTGTGTGAGA 1088
Db 374 ATGAGACGGCCGGTGTCTGTGGCTTCAAAATGGCGCTCAGCAACAACTGTGTGTGAGA 315
Qy 1089 TCGCCAGCTTGCGATTCCTTCTTATCAACAGAGACTTATGCTCATCAGGCTG 1148
Db 314 TCGCCAGCTTGCGATTCCTTCTTATCAACAGAGACTTATGCTCATCAGGCTG 255
Qy 1149 AGGCTGCCAAGATCAACACAGTAGAAGTAGTGAGAGTGGGAGTGGGCGCTTCCGCGC 1208
Db 254 AGGCTGCCAAGATCAACACAGTAGAAGTAGTGAGAGTGGGAGTGGGCGCTTCCGCGC 195
Qy 1209 CTGTTTCCCTCCCACTCCAGCCCTTGAAGTGGGAACTCAGGCTCTGGAAGAAACCTG 1268
Db 194 CTGTTTCCCTCCCACTCCAGCCCTTGAAGTGGGAACTCAGGCTCTGGAAGAAACCTG 135
Qy 1269 GGCAGTGGAGAGCTCAAGTGGCGGCAATTTAGAGCTTTGAGGAGATAGGGCTGG 1328
Db 134 GGCAGTGGAGAGCTCAAGTGGCGGCAATTTAGAGCTTTGAGGAGATAGGGCTGG 75
Qy 1329 CTTTGTGAACCGACAGAGCTGAGAACTCAGGCTTCCCTAGATCCAGAGCCCTCC 1388
Db 74 CTTTGTGAACCGACAGAGCTGAGAACTCAGGCTTCCCTAGATCCAGAGCCCTCC 15
Qy 1389 CATCTTCTCTCTC 1402
Db 14 MATCTTCTCTCTC 1

```

```

RESULT 3
AL523921 927 bp mRNA linear EST 13-PEB-2001
LOCUS AL523921 LTI_NFL003.NBC3 Homo sapiens cDNA clone CS0DC003YA08 5
DEFINITION AL523921 mRNA sequence.
ACCESSION AL523921
VERSION AL523921.1 GI:12787414
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS Li, W. B., Gruber, C., Jessup, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

FEATURES

```

source
1..927
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC003YA08"
/clone_lib="LTI_NFL003_NBC3"

```

```

/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 216 a 261 c 268 g 169 t 13 others
ORIGIN

```

```

Query Match 57.3%; Score 890; DB 9; Length 927;
Best Local Similarity 97.1%; Pred. No. 4.5e-156;
Matches 901; Conservative 13; Mismatches 13; Indels 1; Gaps 1;

```

```

Qy 163 TCAGGTGACCAAGATCTGGGCGCTGGAAGCTGCAAAATGACCTCGTCATTAATGC 222
Db 1 TCAGGTGACCAAGATCTGGGCGCTGGAAGCTGCAAAATGACCTCGTCATTAATGC 60
Qy 223 TGTGTGACAGCAACGACCTGACCTCCGTCCACCTGGCGGTGAATTCAGCTGCCCTT 282
Db 61 TGTGTGACAGCAACGACCTGACCTCCGTCCACCTGGCGGTGAATTCAGCTGCCCTT 120
Qy 283 CACCCTTGGGAGTCCAGAGCGCTGTACGCCCTGTCTACAGATCTGTATCTCCA 342
Db 121 CACCCTTGGGAGTCCAGAGCGCTGTGTACGCCCTGTCTACAGATCTGTATCTCCA 180
Qy 343 GTTGGCGCTGTGACCAATGAGGAGCTGCAACCAAGGCTATTGACAGCATCCAGAGCA 402
Db 181 GTTGGCGCTGTGAGG-CATGAGGAGCTGCAACCAAGGCTATTGAGCATCCAGAGCA 239
Qy 403 GGCCCTGTTTGAAGAGCTGAGAGAGCTGCTGAGCAAAATGGGATGCAACGAGCAACC 462
Db 240 GGCCCTGTTTGAAGAGCTGAGAGAGCTGCTGAGCAAAATGGGATGCAACGAGCAACC 299
Qy 463 CACCTTGAAGACTTTCAGAGACTGCTGACAGACACCTGTATGCTTCTTACTGGCCG 522
Db 300 CACCTTGAAGACTTTCAGAGACTGCTGACAGACACCTGTATGCTTCTTACTGGCCG 359
Qy 523 TACCCGCAAGGCGCTGAGAGGCGCACTGGACACTCATGAAGAGATTAACCTGCGAGAGA 582
Db 360 TACCCGCAAGGCGCTGAGAGGCGCACTGGACACTCATGAAGAGATTAACCTGCGAGAGA 419
Qy 583 CCAGACAGTGCAGCGCTGCGCCAAAGGGGACCAAGTGTGAATCTTCTGATGACAGAGA 642
Db 420 CCAGACAGTGCAGCGCTGCGCCAAAGGGGACCAAGTGTGAATCTTCTGATGACAGAGA 479
Qy 643 CCTGATTTGATGACAGTAAAGCTCAAAGACATGCGAGATGAGTCTCTGGAACATGAGCTGAT 702
Db 480 CCTGATTTGATGACAGTAAAGCTCAAAGACATGCGAGATGAGTCTCTGGAACATGAGCTGAT 539
Qy 703 GGTGGCTGACCGGCGCCAGAGCGAAGAAATTAAGGCAAGGAGCAAGAACTGCAATAAGTT 599
Db 540 GGTGGCTGACCGGCGCCAGAGCGAAGAAATTAAGGCAAGGAGCAAGAACTGCAATAAGTT 599
Qy 763 GCAGGTGCTAGTGAACAGATCAACAGGATGAGCTCTCCGAGCTTCGCAACCAAGACACT 822
Db 600 GCAGGTGCTAGTGAACAGATCAACAGGATGAGCTCTCCGAGCTTCGCAACCAAGACACT 659
Qy 823 GGCAGTGTGGGGGCGCGCATGTGCGGTACCTGATGCGCTCGCGTGAATCACCTTGGG 882
Db 660 GGCAGTGTGGGGGCGCGCATGTGCGGTACCTGATGCGGTGCGGTGARAACACCTTGGG 719
Qy 883 CAGAGCAACCAAGATTAACCAAGATTTGATGTGAGACTGTCTTGAAGGCTCGGCGCTGAA 942
Db 720 CAGAGCAACCAAGATTAACCAAGATTTGATGTGAGACTGTCTTGAAGGCTCGGCGCTGAA 779
Qy 943 GATATCCCGAAACAAAGGTGTATCAAGCTGAAGAACAAAGGTATTTCTTCATTGCCAA 1002

```

Db 780 GATATCCGGAAACAAGTGTCTATCACTGAAGAAACAAGGATTTCTTCAATGCCAA 839

Qy 1003 TGAGGCTGACGGCCCATCTACATCGATGACGGCCGGCTCTCTGTGCTCCAAATGAGCG 1062

Db 840 TGAGGCTGACGGCCCATCTACATCGATGACGGCCGGCTCTCTGTGCTCCAAATGAGCG 899

Qy 1063 CCTCAGCAACAATCTCTGTGCTGAAGTC 1090

Db 900 CCTCAACACMACCTCTGTGTGAGATC 927

RESULT 4
AL575639/c 993 bp mRNA linear EST 16-FEB-2001
LOCUS AL575639 LTI_NFL006.PL2 Homo sapiens cDNA clone CS00D1069YL09 3
DEFINITION prime, mRNA sequence.
ACCESSION AL575639
VERSION AL575639.1 GI:12936998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 993)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers

FEATURES
source 1..993
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS00D1069YL09"
/clone_11b="LTI_NFL006.PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 206 a 276 c 285 g 220 t 6 others
ORIGIN

Query Match 56.9%; Score 883.6; DB 9; Length 993;
Best Local Similarity 96.1%; Pred. No. 6.6e-155;
Matches 971; Conservative 4; Mismatches 14; Indels 21; Gaps 6;

Qy 488 CTGCACAGACACCTGATGCTCTTCTACTCGGCGGACCGGAGGCGCTGCAGGCCCA 547
Db 990 CTGTGCAAAACCTGATCTCTTCTACT- GCCGTACCGGAAAG-CTGCAGGCCCA 933

Qy 548 TGGCAGCTCATGAAGAGATTAATCTCTGCTGAGAGACAGACAGAGTGAAGCGCTGCCCAA 607
Db 932 TGGCAGCTCATGAAGAGATTAATCTCTGCTGAGAGACAGAGACAGAGTGAAGCGCTGCCCAA 878

Qy 608 GGGGACCAAGTGTGCTGCTCTGATGACAGAGACCTGATGATGACAGTAAGCTCAAG 667
Db 877 -GGGACCAAGTGTGCTGCTCTGATGACAGAGACCTGATGATGACAGTAAGCTCAAG 819

Qy 668 GACATGCGAATGAGGTCTCTGGAACATGAGCTGATGCTGACCGGCGCCAGAGCGA 727
Db 818 GACATGCGAATGAGGTCTCTGGAACATGAGCTGATGCTGACCGGCGCCAGAGCGA 759

Qy 728 GAGATTGCGAGCTGGAACAGAACTGCAATAGTGGCAGGTCTAGTGAAGACATCACA 787

Db 758 GAGATTGCGAGCTGGAACAGAACTGCAATAGTGGCAGGTCTAGTGAAGACATCACA 699

Qy 788 GGCATGAGCTCTCCGGAATCTTGACAACAAGACATGCGAGTCTCGGGCGGCTAGTGT 847

Db 698 GGCATGAGCTCTCCGGAATCTTGACAACAAGACATGCGAGTCTCGGGCGGCTAGTGT 639

Qy 848 CGGTACCTGATGAGGTCTCGGTGAGATCAACCTGGGACAGCAACCAAGATTAACAGATT 907

Db 638 CGGTACCTGATGAGGTCTCGGTGAGATCAACCTGGGACAGCAACCAAGATTAACAGATT 579

Qy 908 GATGTGACCTGCTCTGTGAGGGTCCGGCTGGAAGATATCCCGAAACAAGTGTCAATC 967

Db 578 GATGTGACCTGCTCTGTGAGGGTCCGGCTGGAAGATATCCCGAAACAAGTGTCAATC 519

Qy 968 AGCTGGAAGACACGAGTATTTCTTCAATGCGCAATGAGGGTGCAGCGCCATCTACATC 1027

Db 518 AGCTGGAAGACACGAGTATTTCTTCAATGCGCAATGAGGGTGCAGCGCCATCTACATC 459

Qy 1028 GATGAGCGGCGGTCTCTGTGAGCTCCAAATGGGCGCTCAGCAACAATCTGTGTGAG 1087

Db 458 GATGAGCGGCGGTCTCTGTGAGCTCCAAATGGGCGCTCAGCAACAATCTGTGTGAG 399

Qy 1088 ATGCGACGCTCGATGCTGCTCTCTTATCAACAGACCTGATGCTCTCATCAGGGCT 1147

Db 398 ATGCGACGCTCGATGCTGCTCTCTTATCAACAGACCTGATGCTCTCATCAGGGCT 339

Qy 1148 GAGGCTGCCAAGATCAACACAGATGAGAGATGTGCGAGACTGTGGGCCCTCTCCG 1207

Db 338 GAGGCTGCCAAGATCAACACAGATGAGAGATGTGCGAGACTGTGGGCCCTCTCCG 279

Qy 1208 CTTGTTTCCCTGCAACCTCCAGCTTGTGAGCTGAGAGCTGAGAGAGCTTGTGAGAGAG 1267

Db 278 CTTGTTTCCCTGCAACCTCCAGCTTGTGAGCTGAGAGCTGAGAGAGCTTGTGAGAGAG 219

Qy 1268 GGGCAGTGGAGGCTCAGCTGCGGCGCATTTGATTTGAGCTTTGAGAGAGATGAGGCTG 1327

Db 218 GGGCAGTGGAGGCTCAGCTGCGGCGCATTTGATTTGAGAGCTTTGAGAGAGATGAGGCTG 159

Qy 1328 GCTTTTGTGAGACAGAGAGGCTGAGAACTCAGAGCTTCTTATGATCCAGAGCCCTCC 1387

Db 158 GCTTTTGTGAGACAGAGAGGCTGAGAACTCAGAGCTTCTTATGATCCAGAGCCCTCC 99

Qy 1388 CATCTTCTCTCTTAAACCAACCTACCCCATCTACCCCATTTGACACCTTCA 1447

Db 98 CATCTTCTCTCTTAAACCAACCTACCCCATCTACCCCATTTGACACCTTCA 51

Qy 1448 CTCCTGTGCTCAGCTGA-TTGGCTCAGACTCTTCTTTATGTTTTT 1496

Db 50 CTCCTGTGCTCAGCTGANTTAGCTCAGACTCTTCTTTATGTTTTT 1

RESULT 5
AL576040/c 913 bp mRNA linear EST 16-FEB-2001
LOCUS AL576040 LTI_NFL006.PL2 Homo sapiens cDNA clone CS00D1068YA02 3
DEFINITION prime, mRNA sequence.
ACCESSION AL576040
VERSION AL576040.1 GI:12937792
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 913)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers

FEATURES


```

source
1. 913
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="CS0D1068YA02"
/clone_lib="LTI_NFL006_P12"
/issue_type="Placenta"
/notice="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      190 a      257 c      260 g      205 t      1 others
ORIGIN

Query Match      56.4%; Score 876.4; DB 9; Length 913;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 910; Conservative 1; Mismatches 2; Indels 13; Gaps 2;

QY 566 TATTACCTGCTGAGAGACCAAGTCAGCCGCTGCCAAAGGAGGACCAAGTGTGAAC 625
DB 913 TATTACCTGCTGAGAGACCAAGTCAGCCGCTGCCAAAGGAGGACCAAGTGTGAAC 855
QY 626 TTCTCTGATGAGAGAGACCTGATTGATGACATTAAGCTCAAGACATGCGAGATGAGTTC 685
DB 854 TTCTCTGATGAGAGAGACCTGATTGATGACATTAAGCTCAAGACATGCGAGATGAGTTC 795
QY 686 CTGAACATGAGCTGATGATGCTGACCGGCGCCAGAAACGAGAGATTGCGGAGCTGGAA 745
DB 794 CTGAACATGAGCTGATGATGCTGACCGGCGCCAGAAACGAGAGATTGCGGAGCTGGAA 735
QY 746 CAGGAATGACATTAAGTGGCAGGTGCTAGTGAACAGATCAGAGCATGAGCTCTCCGAC 805
DB 724 CAGGAATGACATTAAGTGGCAGGTGCTAGTGAACAGATCAGAGCATGAGCTCTCCGAC 675
QY 806 TTTCAGCAACACACACATGCGCACTGCTGCGGGCCGATGTCGGTACTGATGCGCTTG 865
DB 674 TTTCAGCAACACACACATGCGCACTGCTGCGGGCCGATGTCGGTACTGATGCGCTTG 615
QY 866 CGTGAGATCAACCTGCGGAGAGCAACCAAGATTAACAGATTGATGAGCTGTCTCTG 925
DB 614 CGTGAGATCAACCTGCGGAGAGCAACCAAGATTAACAGATTGATGAGCTGTCTCTG 555
QY 926 GAGGCTCCGCGCTGGAAGATATCCCGGAAACAAGGTGTCATCAAGCTGAAGAACAACGT 985
DB 554 GAGGCTCCGCGCTGGAAGATATCCCGGAAACAAGGTGTCATCAAGCTGAAGAACAACGT 495
QY 986 GATTTCCTTCAATGCAATGAGGTGACCGCCCATCTTACATGATGAGACGCGCGTCTC 1045
DB 494 GATTTCCTTCAATGCAATGAGGTGACCGCCCATCTTACATGATGAGACGCGCGTCTC 435
QY 1046 TGTGAGCTCAAAATGGGCGCTCAGCAACCACTGCTGAGTGGAGATGCGCCGCTGGATTTC 1105
DB 434 TGTGAGCTCAAAATGGGCGCTCAGCAACCACTGCTGAGTGGAGATGCGCCGCTGGATTTC 375
QY 1106 GTCTTCCTTATCAACACAGAGACCTCAATGGCTCATCAGAGGCTGAGGCTGCAAGATCA 1165
DB 374 GTCTTCCTTATCAACACAGAGACCTCAATGGCTCATCAGAGGCTGAGGCTGCAAGATCA 315
QY 1166 CCAAGTGAAGATGATGAGGAGACCTGCGGCGCTCTCTCCGCGCTGTTTCCCTGCGACT 1225
DB 314 CCAAGTGAAGATGATGAGGAGACCTGCGGCGCTCTCTCCGCGCTGTTTCCCTGCGACT 255
QY 1226 CAGAGCCCTTGAAGCTGAGGAGACCTGAGGAGAAACCTGAGGAGAGGAGGAGTGAC 1285
DB 254 CAGAGCCCTTGAAGCTGAGGAGACCTGAGGAGAAACCTGAGGAGAGGAGGAGTGAC 195
QY 1286 CTGCGGCGCATTTGATGAGCTTTGAGGAGAGATGAGGCTGCTTTGTGAAGCAGAC 1345

```

```

DB 194 CTGCGGCGCATTTGATGAGCTTTGAGGAGAGATAGGCGTGGCTTTGTGAAGCCAGAC 135
QY 1346 GAGGCTGAGAACCTCAGGCTTCCCTAGAACAGAGCCCTCCCATCTTCTCTCTAA 1405
DB 134 GAGGCTGAGAACCTCAGGCTTCCCTAGAACAGAGCCCTCCCATCTTCTCTCTAA 75
QY 1406 AAACAACCTTACCCCATTTACCCCATTTGACACCTTCACTCTGTCTCCAGCTG 1465
DB 74 AAACAACCTTACCCCATTTGACACCTTCACTCTGTCTCCAGCTG 27
QY 1466 ATTACCTTCAAGACTTCTTTTATG 1491
DB 26 ATTACCTTCAAGACTTCTTTTATG 1

RESULT 6
AL581415/c 1000 bp mRNA linear EST 16-FEB-2001
LOCUS AL581415 LTI_FL011_BCl Homo sapiens cDNA clone CS0DG003Y101 3 prime
DEFINITION , mRNA sequence.
ACCESSION AL581415
VERSION AL581415.1 GI:12948392
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayres, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
BP 191 91006 Evry cedex - France
Email: seqrel@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
source
1. 1000
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="CS0DG003Y101"
/clone_lib="LTI_FL011_BCl"
/sex="male"
/issue_type="B cells from Burkitt lymphoma"
/lab host="DH10B"
/notice="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      218 a      278 c      282 g      214 t      8 others
ORIGIN

Query Match      56.1%; Score 872; DB 9; Length 1000;
Best Local Similarity 94.1%; Pred. No. 9.5e-153;
Matches 947; Conservative 6; Mismatches 36; Indels 17; Gaps 4;

QY 500 CCTGATGCTTCTTACCTGCGCGTACCGGAAGCCCTGACAGGCGCATGCGAGCTCATG 559
DB 990 CCTTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 934
QY 560 AAGCAGATTAACCTGCTGAGAGACCAAGTCAGCCGCTGCGCAAGGAGGAGCAAGTG 619
DB 934 AAGCAGATTAACCTGCTGAGAGACCAAGTCAGCCGCTGCGCAAGGAGGAGCAAGTG 874
QY 620 CTGAACCTTCTGATGAGAGAGACCTGATTGATGACATTAAGCTCAAGACATGAGAGAT 679
DB 873 CTGAACCTTCTGATGAGAGAGACCTGATTGATGACATTAAGCTCAAGACATGAGAGAT 814
QY 680 GAGTCTCGAATGAGCTGATGATGCTGACCGGCGCCAGAAACGAGAGATTGCGCAG 739

```

```

Db 813 GAGGCTCTGGAACATGAGTGTGCTGACCGGCCCAAGAGAGATTCGGCAG 754
Qy 740 CTGGAACAGAACTGCAATAGTGGCAGGTGCTAGTGAACATCAAGCATGAGCTCT 799
Db 753 CTGGAACAGAACTGCAATAGTGGCAGGTGCTAGTGAACATCAAGCATGAGCTCT 694
Qy 800 CCGGAGCTTGAACACAGACATGCGAGTGTGCGGGCGGCAATGATGAGTGAAGT 859
Db 693 CCGGAGCTTGAACACAGACATGCGAGTGTGCGGGCGGCAATGATGAGTGAAGT 634
Qy 860 CCGCTCGCTGAGATCACTCTGGGACAGAGCAACCAAGATTAACAGATGATGAGCTG 919
Db 633 CGCTCGCTGAGATCACTCTGGGACAGAGCAACCAAGATTAACAGATGATGAGCTG 574
Qy 920 TCTCTGAGAGGCTCGGCTCTGGAAGATATCCCGAAACAAAGTGTCTCAAGCTGAAGAAC 979
Db 573 TCTCTGAGAGGCTCGGCTCTGGAAGATATCCCGAAACAAAGTGTCTCAAGCTGAAGAAC 514
Qy 980 AACGGGATTTCTTCAATGCGCAATGAGGGTGCACGGGCCCATCTCAATGATGAGAGCGCG 1039
Db 513 AACGGGATTTCTTCAATGCGCAATGAGGGTGCACGGGCCCATCTCAATGATGAGAGCGCG 454
Qy 1040 GTGCTCTGTGCTCCAAATGAGCGCTCAGACCAACAACTGTGTGAGATGCGCAAGCTG 1099
Db 453 GTGCTCTGTGCTCCAAATGAGCGCTCAGACCAACAACTGTGTGAGATGCGCAAGCTG 394
Qy 1100 CGATTCGTCTTCTTATCAACAGAGCTCTGATGAGGCTGAGGCTGAGCTGCAAG 1159
Db 393 CGATTCGTCTTCTTATCAACAGAGCTCTGATGAGGCTGAGGCTGAGCTGCAAG 334
Qy 1160 ATCAACCAACATGAGGAGATGAGGAGCACTGTGGGCGCTCTCGGCTGTTCCTCCCT 1219
Db 333 ATCAACCAACATGAGGAGATGAGGAGCACTGTGGGCGCTCTCGGCTGTTCCTCCCT 274
Qy 1220 GCACTCCAGCCCTCTGAGTGGAGAACTCAGGCTCTGGAACAAACCTGGAGAGTGGAG 1279
Db 273 GCACTCCAGCCCTCTGAGTGGAGAACTCAGGCTCTGGAACAAACCTGGAGAGTGGAG 214
Qy 1280 GCTCAGCTCGGGGCTCTGATGAGCTCTTGAAGAGAGATGAGGCTGAGCTTGTGAAG 1339
Db 213 GCTCAGCTCGGGGCTCTGATGAGCTCTTGAAGAGAGATGAGGCTGAGCTTGTGAAG 154
Qy 1340 CGAGCAGAGCTGAGAACTCTCAGGCTTCCCTAGATCAAGCCCTCCCATCTCTCT 1399
Db 153 CGAGCAGAGCTGAGAACTCTCAGGCTTCCCTAGATCAAGCCCTCCCATCTCTCT 94
Qy 1400 CTCTAAACCAACCTTACCCCATCTTACCCCATCTTACCTCACTCTGTGTCT 1459
Db 93 CCSSCAACACCC-----CTACCCCTTTCCTCTTACT-CTTTTCTC 48
Qy 1460 CAGCTGATTAAGCTCAGACTCTTCTTATATGTTTTCTTTGTAA 1505
Db 47 CTGCTGATTAATCTTCAGATTTCTTCTTATGTTTTCTTTGTAA 2

```

```

RESULT 7
AL523920/c 1133 bp mRNA linear EST 13-FEB-2001
LOCUS AL523920 LTI_NFL003 NBC3 Homo sapiens cDNA clone CS0DC003YA08 3
DEFINITION prime, mRNA sequence.
ACCESSION AL523920
VERSION AL523920.1 GI:12787413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1133)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polajew, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

```

```

Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    source
        1..1133
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="CS0DC003YA08"
            /clone_lib="LTI_NFL003_NBC3"
            /sex="male"
            /tissue_type="neuroblastoma cells"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@life.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 229 a 308 c 314 g 241 t 41 others
ORIGIN

```

```

Query Match 55.7%; Score 864.6; DB 9; Length 1133;
Best Local Similarity 91.5%; Pred. No. 2.1e-151;
Matches 930; Conservative 32; Mismatches 49; Indels 5; Gaps 5;

Qy 410 TTTCAGAGGCTGAGAGAGCTGTGAGCAAGTGGATGACAGCCAGCCACTTG 469
Db 1024 TTTTGAAGAGTGAAGAGAGCTG-TGAGCAAGTGG-TGACAGCAGCCAGCCACTTG 967
Qy 470 GAGACTTTCAGAGCTGTGTGAGAGACCTGATGCTTACTGCTGAGCCGCTACCGG 529
Db 966 AAGMCTTCCAGAGCCCTGTAAGAGACCTGATGCTTCTGCTGAGCCGCTACCGG 908
Qy 530 AAGGCTGAGAGGCTGAGAGCTGAGAGAGTGAAGAGTGAAGTGAAGAGAGAGAG 589
Db 907 AAGGCTGAGAGGCTGAGAGCTGAGAGAGTGAAGAGTGAAGTGAAGAGAGAGAG 849
Qy 590 GTGACAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649
Db 848 GAGCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 790
Qy 650 GATGACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 709
Db 789 GATTAACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
Qy 710 GACCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 769
Db 729 AACCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
Qy 770 CTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 829
Db 669 CTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610
Qy 830 CTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 889
Db 609 CTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550
Qy 890 ACCAAGAGATTAACAGATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 949
Db 549 ACCAAGAGATTAACAGATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 490
Qy 950 CGGAACAGAGTGTATCAAGCTGAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1009
Db 489 CGGAACAGAGTGTATCAAGCTGAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 430
Qy 1010 CGACGAGCCATTTACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1069
Db 429 CGACGAGCCATTTACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370

```

QY 1070 AACCACTCGTGTGGAGATATCGCAGGCGTGGCATTCGTTCTCTTATCAACGAGACCTC 1129

Db 369 AACCACTCGTGTGGAGATATCGCAGGCGTGGCATTCGTTCTTATCAACGAGACCTC 310

QY 1130 ATTGACCTCATCAGGGCTAGGCTGCGCAAGATCAACCAACAGTGAAGAAATGGTGCGAGA 1189

Db 309 ATTGACCTCTTACAGGGCTAGGCTGCGCAAGTTCACACCAACAGTGAAGAAATGGTGCGAGA 250

QY 1190 CTGCTGGGCCCTCTCCGGCTGTTCCTCCCTGCACCTCAAGCCCTTGAAGCTGGGAATCT 1249

Db 249 CTGCTGGGCCCTCTCCGGCTGTTCCTCCCTGCACCTCAAGCCCTTGGGGCTGGGGACCTC 190

QY 1250 AGGCTTCCTGGAAAAACCTGGGCGAGTGGGAGGCTCAGCTCGGGGCATTGATTTGAGCTTT 1309

Db 189 AGGCTTCCTGGAAAAACCTGGGCGAGTGGGAGGCTCAGCTCGGGGCATTGATTTGAGCTTT 130

QY 1310 TGAAGGAGGATAGGGCTGGGCTTTTGTGAAGCCAGACGAGGCTGAGAACTCAGGCTTCCC 1369

Db 129 TGAAGGAGGATAGGGCTGGGCTTTTGTGAAGCCAGACGAGGCTGAGAACTCAGGCTTCCC 70

QY 1370 TAGATTCAGAGGCCCTCCCATCTTCCTCCTTAATAAACCAACCCATACCCCATTT 1425

Db 69 TAGATTCAGAGGCCCTCCCATCTTCCTCCTTAATAAACCAACCCATACCCCATTT 14

LOCUS	AL554401	882 bp	mRNA	linear	EST 16-FEB-2001
DEFINITION	AL554401 LTI_NFL006_PL2 Homo sapiens CDNA clone CS0D1004YM07 3				

ACCESSION	prime, mRNA sequence.
AL544010	
AL544010.1	GI:12876489

KEYWORDS EST.
SOURCE human.
OBJECTS Homo sapiens

ORGANISM
humo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 882)	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.	Full-length cDNA libraries and normalization

JOURNAL
COMMENT
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES	location/Qualifiers
source	1. .882
	/external-llwms.candidate

```

/orgall.html?noio_sapient-
/db_xref="taxon:9606"
/c1one="CS0DI004YM07"

```

```
/clone_lib="LTI_NFL006_PL2"  
/tissue_type="pitcenta"  
/note="Vector: pCMVSPORT 6: Site 1: NotI: 1st strand cDNA
```

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Jiang Life Technologies

a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : ftjano@lifestar.com ftjano@lifestar.com ftjano@lifestar.com

```

BASE COUNT      175 a      251 c      253 g      199 t      4 others

```

Query Match	Score	DB	Length
54.4%	844.8	9	882

Best Local Similarity	97.6%;	Pred. No. 1.2e-147;			
Matches	868;	Conservative 3;	Mismatches 7;	Indels 11;	Gaps 1;

QY 578 GAGGACCGAGCAGTGCAGCCGCTGCCCAAGGGGACCAAGTGCTGAATTCTGTATGCA 637

Db 878 GAAGACCAGACGTGCACCCGCTGCCCAARGGAGCCAACTTCTCTGATGCA 81.9

QY	638	GAGGACCTGATTGATGAGACAGTAAAGCTCAAGGACATGCGAGATGAGGCTCTGGAAATATAG	697
Db	818	GAGGACCTGATTGATGAGACAGTAAAGCTCAAGGACATGCGAGATGAGGCTCTGGAAATATAG	759
QY	698	CTGATGCTGGCTGACACCGGCGCCAGAAAGGAGAGATTTCCGACAGCTGGAAACAGGAATCTGCAT	757
Db	758	CTGATGCTGGCTGACACCGGCGCCAGAAAGGAGAGATTTCCGACAGCTGGAAACAGGAATCTGCAT	699
QY	758	AAAGTGCAGGTGCTTAGTGAACAGCATCAACAGCATAGACTCTCCGGACTTCGACAAACAG	817
Db	698	AAAGTGCAGGTGCTTAGTGAACAGCATCAACAGCATAGACTCTCCGGACTTCGACAAACAG	639
QY	818	ACACTGGCAGTCTCTCGGCGGCGCCAGTGGTGCGGTAACCTGATCGCTCCGCTGAGATCAC	877
Db	638	ACACTGGCAGTCTCTCGGCGGCGCCAGTGGTGCGGTAACCTGATCGCTCCGCTGAGATCAC	579
QY	878	CTGGGCAAGCAACCAAGATTAACAGATTGATGTGACCTGTCTCTGAGAGGTCGGGCC	937
Db	578	CTGGGCAAGCAACCAAGATTAACAGATTGATGTGACCTGTCTCTGAGAGGTCGGGCC	519
QY	938	TGGAGATATCCCGGAAACAAGGTCTCAATCAAGTGAAGAACAAACGGTATTTCTTCATT	997
Db	518	TGGAGATATCCCGGAAACAAGGTCTCAATCAAGTGAAGAACAAACGGTATTTCTTCATT	459

998 GCCCATGAGGCTCGACGGCCCATCTACATGATGAGCGCGGCTCTGTGGCTCCAA 1057

Db 458 GCCAATGAGGGTCGACGGCCCATCTACATCGATGGACGGCCGGTCTCTGTGGCTCCAA 399

Db 398 TGGGCGCTCAGCAACACTCTGTGTGAGATGCCAGCCTCGATTGCTTCCATTATC 339

1118 AACGAGACCTCATTGCCCCTCATCAGGGCTGAGGCTGCCAAGATCACACCAAGTGAAGA 1177

Db 338 AACGAGACCTCATGGCCCTCATGAGGCTGAGGCTGCCAAGATCACACCAACAGTAGGA 279

Qy 11/8 ATGGGCAAGACATCGTGGCCCTCTCGGCGCTGTTCCCTCGACATCCAGGCCCTTG 1237

QY 1238 AGCTGGAACTCAGGCTCCTGGAAAACTGGGCAGTGGGAGGCTCAGCTGGGGCCATT 1297

Db 218 AGCTGGGAATCAGGCTCCTGSAANAACCTGGGCAGTGGGAGGCTCAGCTGGCGGCGCATT 159

QY 1298 GATTTGAGCCTTTGAGGGAGATAGGCGTGCCCTTTGTGAAGCAGCAGGGCTGAGAAC 1357

20 1358 CTCAGGCTTCCCTAGATCCAGAGGCCCTCCCACTTTCCTCTCTTAAAAACAACCTAC 1417

Db 98 CTCAGGCTTCCCTAGATCCAGAGCCCCCTCCCACATCTTCTCTCTTAATAACAACGC --- 42

QY 1418 CCCCATTCACCCCCATTTGCGACCTTCACCTCTGTGTCTCCAGCTGA 1466

```

UD      41 -----C1CCCCCCC1GCCACC1NAC1CC1G1G1C1C1CCAG1GA I

```

RESULT 9	
BMS49097	
LOCUS	1021 bp
BMS49097	linear
	EST 20-FEB-2002

DEFINITION	AGNCOUNT	NIH_MGC_72	Homo sapiens	cdna clone	IMAGE:5548420
5', mRNA sequence.					

ACCESSION BM549097
 VERSION BM549097.1
 KEYWORDS GI:18784222
 EST.

SOURCE ORGANISM
human.
Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

REFERENCE

1 (bases 1 to 1021)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM1257 row: F column: 05
High quality sequence stop: 731.
Location/Qualifiers
1..1021

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5548420"
/clone_lib="NIR MGC 72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 232 a 292 c 292 g 195 t 10 others

ORIGIN

Query Match 53.9%; Score 836.6; DB 13; Length 1021;
Best Local Similarity 96.8%; Pred. No. 3.7e-146;
Matches 907; Conservative 0; Mismatches 21; Indels 9; Gaps 5;

151 TAAACAGCAGCATTGAGTACCAAGATCTGGGCGCTGGAAGCTGCAATGACCTCT 210
Db 1 TAAACAGCAGCATTGAGTACCAAGATCTGGGCGCTGGAAGCTGCAATGACCTCT 60

211 GCTCAATAATGCTGTGTTCAGACCAACGACCTGACCTCGTCCACCTGGGCGTGAATT 270
Db 61 GCTCAATAATGCTGTGTTCAGACCAACGACCTGACCTCGTCCACCTGGGCGTGAATT 120

271 CAGCTGCCCTTACCTTCGGGAGGTCCAGAGAGCTTGTACGCTTCTCTACGATCC 330
Db 121 CAGCTGCCCTTACCTTCGGGAGGTCCAGAGAGCTTGTACGCTTCTCTACGATCC 180

331 TGTCAATCTCAAGTTGGCTGACGGCCATGAGGAGCTGACCCGAGAGCTATTGACG 390
Db 181 TGTCAATCTCAAGTTGGCTGACGGCCATGAGGAGCTGACCCGAGAGCTATTGACG 240

391 CATCCAGACCAAGGCGCTTTAGCAAGGCTGAGAGAGCTGCTGAGCAAAAGTGGATC 450
Db 241 CATCCAGACCAAGGCGCTTTAGCAAGGCTGAGAGAGCTGCTGAGCAAAAGTGGATC 300

451 GACCAAGCCGACCTTGGAGACCTTCCAGAGACCTGCTGCACAGACACCTTGATGCC 510
Db 301 GACCAAGCCGACCTTGGAGACCTTCCAGAGACCTGCTGCACAGACACCTTGATGCC 360

511 CTAACCGGCGCGTACCGGAGGCGCTGAGGCGCACTGAGAGCTCATGAAGCAGATT 570
Db 361 CTAACCGGCGCGTACCGGAGGCGCTGAGGCGCACTGAGAGCTCATGAAGCAGATT 420

571 CCGTGTGAGGACCAAGACAGTGCAGCGCTGCCCAAGGGGACCAAGTCTGAATCTTC 630
Db 421 CCGTGTGAGGACCAAGACAGTGCAGCGCTGCCCAAGGGGACCAAGTCTGAATCTTC 480

631 TGAATGACAGAGACCTGATTGATGACAGTAAGCTCAAGAGACATGCGAGATGAGTCTTGA 690
Db 481 TGAATGACAGAGACCTGATTGATGACAGTAAGCTCAAGAGACATGCGAGATGAGTCTTGA 540

691 AATATGAGTGAATGAGTGGTGAACCGGCGCAAGAGAGAGATTCGACACTGGAAACGA 750
Db 541 AATATGAGTGAATGAGTGGTGAACCGGCGCAAGAGAGAGATTCGACACTGGAAACGA 600

751 ACTGCATTAATGAGTGGTGAATGAGTGAACAGATCAGAGGATAGTCTCCGAGCTTCCA 810
Db 601 ACTGCATTAATGAGTGGTGAATGAGTGAACAGATCAGAGGATAGTCTCCGAGCTTCCA 660

811 CAACAGACACTGGAGAGTGTGCGGGGCGCATGCTGAGTACCGCTCCGCTGA 870
Db 661 CAACAGACACTGGAGAGTGTGCGGGGCGCATGCTGAGTACCGCTCCGCTGA 720

871 GATCACTGAGGAGAGGACCAACCAAGATTAACCAATGATGAGACTGTCTTGGAGG 930
Db 721 GATCACTGAGGAGAGGACCAACCAAGATTAACCAATGATGAGACTGTCTTGGAGG 780

931 TCCGGCTGAGAGATATCCC-GGAACAAGTGTATCAAGTGA-AAGAACACAGGTAT 988
Db 781 TCCGGCTGAGAGATATCCNNNGAACAGAGTGTATCAAGTGAACACAGGTAT 840

989 TTCTTCATTGGCCATGA-GGGTCAGGCGCCCATCTACA-TCGATGAGAGGCC-----CGT 1041
Db 841 TTCTTCATTGGCCATGAAGAGGTCGACGGCCCATCTACANTGATGAGAGCGCGGTGCT 900

1042 GCTGTGGCTCCAAATGGCGCTGAGGACCAACTCT 1078
Db 901 CCGTGGCTGCCCAATGGCGCTGAGGACCAACT 937

RESULT 10
AL544052 848 bp mRNA linear EST 16-FEB-2001
LOCUS AL544052 LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1004YM07 5
DEFINITION prime, mRNA sequence.
ACCESSION AL544052
VERSION AL544052.1 GI:12876531
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 848)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..848
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1004YM07"
/clone_lib="LTI.NFL006.PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 202 a 235 c 252 g 157 t 2 others

ORIGIN

Query Match 53.6%; Score 833; DB 9; Length 848;
Best Local Similarity 99.4%; Pred. No. 1.9e-145;
Matches 844; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

333 TCATCTCCAAAGTTGGCTGTGACGACATGAGGACCTGACCAAGGCTATTGAGCA 392
Db 1 TCATCTCCAAAGTTGGCTGTGACGACATGAGGACCTGACCAAGGCTATTGAGCA 60

393 TCAGAGCAAGGCGCTGTTTGAAGGCTGAGAGCAAGCTGCTGAGCAAGTGGATCGA 452
Db 61 TCAGAGCAAGGCGCTGTTTGAAGGCTGAGAGCAAGCTGCTGAGCAAGTGGATCGA 120

QY 453 CCAGCCAGCCACCTTGGAGACCTTCCAGACCTGCTCAGACACCCCTGATGCTTCT 512
DB 121 CCAGCCAGCCACCTTGGAGACCTTCCAGACCTGCTCAGACACCCCTGATGCTTCT 180
QY 513 ACCTGCGCCGTTACCGCGAAGCCCTGACAGGCCCTGACCTGACCTGATGAGACGATTTACC 572
DB 181 ACCTGCGCCGTTACCGCGAAGCCCTGACAGGCCCTGACCTGACCTGATGAGACGATTTACC 240
QY 573 TGCTGAGAGACCAAGCACTGACCGCTGCCCAAGGGGACCAAGTGTGAACTTCTCTG 632
DB 241 TGCTGAGAGACCAAGCACTGACCGCTGCCCAAGGGGACCAAGTGTGAACTTCTCTG 300
QY 633 ATGACAGAGACCTGATTTGATGACAGTAACTCAAGACATGCGAGATGAGCTCTGAAAC 692
DB 301 ATGACAGAGACCTGATTTGATGACAGTAACTCAAGACATGCGAGATGAGCTCTGAAAC 360
QY 693 ATGAGCTGATGCTGCTGACCGCGCCAGAAAGCGAGAGATTCGGACGCTGGAACAGAAC 752
DB 361 ATGAGCTGATGCTGCTGACCGCGCCAGAAAGCGAGAGATTCGGACGCTGGAACAGAAC 420
QY 753 TGCTAATGTCGAGGTGCTGATGAGACAGATTCACAGGATGAGTCTCTCCGATCTTCAGCA 812
DB 421 TGCTAATGTCGAGGTGCTGATGAGACAGATTCACAGGATGAGTCTCTCCGATCTTCAGCA 480
QY 813 ACCAGACACTGCGAGTGTGCGCGGCGCGATGCTGCGGATGCTGAGCTGCGCTGAGA 872
DB 481 ACCAGACACTGCGAGTGTGCGCGGCGCGATGCTGCGGATGCTGAGCTGCGCTGAGA 540
QY 873 TCACCTGCGGAGAGCAACCAAGATTAACCAAGATTTATGTCGACCTGCTCTGAGAGGCTC 932
DB 541 TCACCTGCGGAGAGCAACCAAGATTAACCAAGATTTATGTCGACCTGCTCTGAGAGGCTC 600
QY 933 CGGCTGGAAGATATCCCGGAAACCAAGTGTATCAAGCTGAAGAACCAACGGTGAATTTCT 992
DB 601 CGGCTGGAAGATATCCCGGAAACCAAGTGTATCAAGCTGAAGAACCAACGGTGAATTTCT 660
QY 993 TCATTGCAATGAGGGTGAAGCGGCCCATCTATCATGATGAGAGCGCGGGTCTGCTGAGCT 1052
DB 661 TCATTGCAATGAGGGTGAAGCGGCCCATCTATCATGATGAGAGCGCGGGTCTGCTGAGCT 720
QY 1053 CCAATGAGCGCTCAGCAACCACTGTGTGTGAGATGCGCAGCTGCGATTCGTCTTCC 1112
DB 721 CCAATGAGCGCTCAGCAACCACTGTGTGTGAGATGCGCAGCTGCGATTCGTCTTCC 779
QY 1113 TTATCAACCAAGACCTCATTTGCTTCATCAGGGCTGAGGCTGCCAAGATCACACCAAGT 1172
DB 780 TTATCAACCAAGACCTCATTTGCTTCATCAGGGCTGAGGCTGCCAAGATCACACCAAGT 839
QY 1173 GAGGAATGG 1181
DB 840 GAGGATGG 848

RESULT 11
AL527295/c 947 bp mRNA linear EST 13-FEB-2001
LOCUS AL527295 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC021YD22 3
DEFINITION prime, mRNA sequence.
ACCESSION AL527295
VERSION AL527295.1 GI:12790788
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 947)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
CONTACT Genoscope - Centre National de Sequencage
COMMENT BP 191 91006 EVRY cedex - France

FEATURES
source
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..947
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC021YD22"
/clone_1ib="LTI_NFL003_NBC3"
/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Peng Liang life Technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 186 a 267 c 273 g 214 t 7 others
ORIGIN
Query Match 53.3%; Score 827.4; DB 9; Length 947;
Best Local Similarity 96.8%; Pred. No. 2e-144;
Matches 884; Conservative 6; Mismatches 8; Indels 15; Gaps 4;
QY 561 AGCATATTACTCTGCTGGAGACCAAGACAGTCCGCTCCCAAGAGGACCAAGTGC 620
DB 900 AGCATATTACTCTGCTGGA -ACCAAGACAGTCCGCTCCCAAGAGTGC 843
QY 621 TGAATCTCTGATGACAGAGACCTGATTTGACAGTAACTCAAGGACATGCGAGATG 680
DB 842 TGAATCTCTGATGACAGAGACCTGATTTGACAGTAACTCAAGGACATGCGAGATG 783
QY 681 AGGTCTGGAACATGAGCTGATGTGCTGAACCGCGCCAGAAAGCGAGATTCGGACGC 740
DB 782 AGGTCTGGAACATGAGCTGATGTGCTGAACCGCGCCAGAAAGCGAGATTCGGACGC 723
QY 741 TGAACAGGAACCTGATTAAGTGGAGGCTGATGAGACAGATCAACAGGATGAGCTTC 800
DB 722 TGAACAGGAACCTGATTAAGTGGAGGCTGATGAGACAGATCAACAGGATGAGCTTC 663
QY 801 CGGACTTGAACACAGACACTGGCAGTGGCGGGCGCATGGTGGATGACCTGATGC 860
DB 662 CGGACTTGAACACAGACACTGGCAGTGGCGGGCGCATGGTGGATGACCTGATGC 603
QY 861 GCTGCGTGAATGACCTCTGGGAGAGCAACCAAGATTAACCAAGTGTGAGCTGT 920
DB 602 GCTGCGTGAATGACCTCTGGGAGAGCAACCAAGATTAACCAAGTGTGAGCTGT 543
QY 921 CTGTGAGGGTCCGGCTGGAAGATATCCGGAAACCAAGTGTGATCAAGCTGAAGAAC 980
DB 542 CTGTGAGGGTCCGGCTGGAAGATATCCGGAAACCAAGTGTGATCAAGCTGAAGAAC 483
QY 981 ACGGTGATTTCTTATGTCGCAATGAGGGTGAAGCGGCCCATCTAATGATGAGCGCGG 1040
DB 482 ACGGTGATTTCTTATGTCGCAATGAGGGTGAAGCGGCCCATCTAATGATGAGCGCGG 423
QY 1041 TGCTCTGTGCTCAAAATGCGCTCAGCAACCACTGTGTGTGAGATTCGACGCTGC 1100
DB 422 TGCTCTGTGCTCAAAATGCGCTCAGCAACCACTGTGTGTGAGATTCGACGCTGC 363
QY 1101 GATTGCTTTTCTTATCAACAGACCTCATTTGCTTCATCAGGGCTGAGGCTGCAAGA 1160
DB 362 GATTGCTTTTCTTATCAACAGACCTCATTTGCTTCATCAGGGCTGAGGCTGCAAGA 303
QY 1161 TCACACCAAGTGAAGATGATGAGGACCTGAGGCGCTCTCGGCTGATTCCTCCCTG 1220
DB 302 TCACACCAAGTGAAGATGATGAGGACCTGAGGCGCTCTCGGCTGATTCCTCCCTG 243
QY 1221 CCACTCAAGCCCTTGAAGTGGAACTCAGGCTCTTGAAGAAACCTTGGGAGTGGAGG 1280

Db	242	CCACTCCAGCCCCCTTGAAGCTGGAAAC	TGAGGCTCTCTGGAAAAA	CCTGGGAGTGGGAGS	183
Qy	1281	CTCAGCTGGGGGCCATTGATTTTGAGCTTTGAGGAGAGATAGGGCTGGCTTTGTGAAC			1340
Db	182	CTCAGCTGGGGGCCATTGATTTTGAGCTTTGAGGAGAGATAGGGCTGGCTTTGTGAAC			123
Qy	1341	CAGCAGAGGCTAGAACTCAGAGCTTCCCTAG--ATCCAGAGCCCCCTCCCATCTTCTC			1398
Db	122	CAGCAGAGGCTAGAACTCAGAGCTTCCCGAATGSSAGACCCCTCCCATCTACTC			63
Qy	1399	TCTCTAAAAACAACCTTACCCCCCATTTCAACCCCATTTGCCACTTCACTCTGTGTCT			1458
Db	62	TCTCTAAAAACAACCTTAC-----CGCCCATTTGCCACATCACTCTGTGTCT			14
Qy	1459	CCAGCTGATTAGC	1471		
Db	13	CGAGCTGATTAGC	1		

RESULT 12	AL529425	852 bp	mRNA	linear	EST 13-FEB-2001
LOCUS	AL529425/c				
DEFINITION	AL529425 LTI NFKB1 NBCC4 Homo sapiens cDNA clone CS0DD006YA05 3				
ACCESSION	AL529425				
VERSION	AL529425.1				
KEYWORDS	EST.				
SOURCE	human.				

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1 (bases 1 to 852)					
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.					
Full-length cDNA libraries and normalization					
Unpublished (2001)					
Contact: Genoscope					

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

FEATURES	Location/Qualifiers
source	1. .852

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODD006YA05"
/clone_lib="UT1_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetech.com URL : http://fulllength.invitrogen.com"

```

BASE COUNT	163 a	262 c	230 g	180 t	17 others
ORIGIN					

Query Match	53.0%	Score 823.4	DB 9	Length 852
Best Local Similarity	96.6%	Pred. No. 1.2e-143		
Matches 823; Conservative	16; Mismatches	13;	Indels	0; Gaps

Qy	492	ACAGACACCCCTGATGCTCTTACCTCGAGCCCGTACCGCAAGGCCCTGAGAGCCACACCTGGC	551
Db	852	ACARACCCCTGATGCTCTTACCTCGAGCCCGTACCGCAAGGCCCTGAGAGCCACACCTGGC	792
Qy	552	AGCTATGATGAGCAGTTTAACTGCTGAGAGGACACAGACGTGACCGCTGCTGCCCAAAGGGG	6111
Db	792	MGCTATGATGAGCAGTTTAACTGCTGAGAGGACACAGACGTGACCGCTGCTGCCCAAAGGGG	733

QY	612	ACCAAGCTGTAACTTCTCTGATGACAGAGCACTGATGATGACAGTAAGCTCAAGGACA	671
Db	732	ACCGAGTGTGAACCTTCTCTGATGACAGAGCACTGATGATGACAGTAAGCTCAAGGACA	673
QY	672	TTCGAGATGAGGTCTGTGAAACATGAGCTGATGTTGTTGAGTCCGCGCCGAAAGCAGAGCA	731
Db	672	TTCGAGATGAGGTCTGTGAAACATGAGCTGATGTTGTTGAGTCCGCGCCGAAAGCAGAGCA	613
QY	732	TTTCGACGCTGGAAACAGGAACATGCACTTAAGTGGCAGGTGCTAATGGAACAGATCAAGGCA	791
Db	612	TTTCGACGCTGGCAACAGGACATGCACTTAAGTGGCAGGTGCTAATGGAACAGATCAAGGCA	553
QY	792	TGAGCTCTCCGGAATTTCGACAAACAGACACTGGCAGTGTGCGGGGACCGCATGTTGCGT	851
Db	552	TGAGCTCTCCGGAATTTCGACAAACAGACACTGTCAGTGTGCGGGGACCGCATGTTGCGT	493
QY	852	ACCTGATGCGCTCGCGTGAATCACCTTGCGGAGAGCAACCAAGATTAACAGATTGATG	911
Db	492	RCGTATGAGGCTCGCGTGAATCACCTTGCGGAGAGCAACCAAGATTAACAGATTGATG	433
QY	912	TGGACCTGTCTTGAAGGGTCCGGGCTTGGAAATATTCGCCGAACAAGGTGTATCAAGC	971
Db	432	TGGACCTGTCTTGAAGGGTCCGGGCTTGGAAATATTCGCCGAACAAGGTGTATCAAGC	373
QY	972	TGAAGAACAGGGTATTTCTTCATTTGCCAATGAGAGTTCGACGGCCCATCTCATCGATG	1031
Db	372	TGAAGAACACGGTATTTCTTCATTTGCCAATGAGAGTTCGACGGCCCATCTCATCGATG	313
QY	1032	GACGGCCGGTGTCTGTGTGCTCCAAATGAGCGCTCAGCAACAACCTGTGTGTGAGATCG	1091
Db	312	GACGGCCGGTGTCTGTGTGCTCCAAATGAGCGCTCAGCAACAACCTGTGTGTGAGATCG	253
QY	1092	CCAGCTGTGATTTGCTCTTCTTTATCAACCAAGACCTCATTTGCCCTCATACGGGCTGAGG	1151
Db	252	CCRGCTGTGATTTGCTCTTCTTTATCAACCAAGACCTCATTTGCCCTCATACGGGCTGAGG	193
QY	1152	CTGCCAAAGTCAACACACAGTAGAGGATGTTGTGACAGACTCGTGGAGCCCTCTCCGGCTG	1211
Db	192	CTGCCAAAGTCAACACACAGTAGAGGATGTTGTGACAGACTCGTGGAGCCCTCTCCGGCTG	133
QY	1212	TTTTCCCTTGCCACTCCAGCCGCCCTTGAGCTGAGGAATCAAGGCTCTCTGAAAAAATCTGGGC	1271
Db	132	TTTTCCCTTGCCACBCCAGCCGCCCTTTAGCTGAGGAATCAAGGCTCTCTGAAAAAATCTGGGC	73
QY	1272	AGTGGAGGCTTAGCTGGGGGCAATTGATTTAGGCTTTGAAGGAGGATAGGGCTGGCCT	1331
Db	72	AGTGGAGGCTTAGCTGGGGGCAATTGATTTAGGCTTTGAAGGAGGATAGGGCTGGCCT	13
QY	1332	TTTGTGAAGCAG 1343	
Db	12	TTTGTGAAGCAG 1	

RESULT 13	AL571554/c	916 bp	mrna	linear	EST 16-FEB-2001
LOCUS	AL571554				
DEFINITION	AL571554 LTI_NFL006.PL2 Homo sapiens cdna clone CS001024YD06 3 prime, mRNA sequence.				

ACCESSION	AL571554	GI:12928964
VERSION	AL571554.1	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Euarchontariida; Primates; Homo	
	sapiens	

REFERENCE Mammalia; Eumetaria; Filicoides; Cataglyphis; Homoptera; Hemiptera
AUTHORS 1 (bases 1 to 916)
TITLE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.


```
Db 775 ACCGGGGCCAGAGGAGATTCCGCACTGGAAACAGAACTGCAATKAGGCAKTC 716
Qy 771 TAGTGAACGACATCACAGGATGAGCTCTCCGGAATTCCAGAA--CAGACACTGGCAGTG 829
Db 715 TMTGTGACGACATCAACAGGACATGAGCTCTCCGACTCTCCAGCAATCCAGACACTGGCAGTG 656
Qy 830 CCGCGGGGGCCGATGGTGTGGGATACCTGATGCGCTCGCGTGAAGTCAACCTGGGCAAGAGA 889
Db 655 CTGGGGGG--CGCATGGTGGGATKATKATGCGTTCCGCTGAATCAACCTGGGTAAGAGA 597
Qy 890 ACCAAGATTAACAGATTGATGATGGAACCTGTCTCTGGAAGGTCGGGCTGGA--AGATATC 948
Db 596 ACYAAAGATTAACAGATTGATGATGGAACCTGTCTCTGGAAGGTCGGGCTGGAATGATATC 537
Qy 949 CCGGAAACAGAGTGTCTATCAAGCTGAAGAACACGCGTATTTCTTATTGCCAATAGGG 1008
Db 536 CCGGAAATTAAGGTGTCTATCAAGCTGAAGAACACGKTGATTTCTTATTGCCAATAGGG 477
Qy 1009 TCGAGCGGCCATCTACATGATGAGAGCGCGGTCCTGTGGGTCCAAATGGGCGCTCAG 1068
Db 476 TCGAGCGGCCATCTACATGATGAGAGCGCGGTCCTGTGGGTCCAAATGGGCGCTCAG 417
Qy 1069 CACAACCTGTGTGATGATGAGAGCGCGGTCCTGTGGGTCCAAATGGGCGCTCAG 1128
Db 416 CACAACCTGTGTGATGATGAGAGCGCGGTCCTGTGGGTCCAAATGGGCGCTCAG 357
Qy 1129 CATTGCGCTCATGAGGCTGAGGCTGCCAAGATCACACACAGTGAAGATGTGGCAGG 1188
Db 356 CATTGCGCTCATGAGGCTGAGGCTGCCAAGATCACACACAGTGAAGATGTGGTGAAG 297
Qy 1189 ACTCGTGGGCGCTCTCGCGGCTGTCTCCCTGCGACATCCAGCGCCCTTGA--GCTGGGAAC 1247
Db 296 ACTGTGGGCGCTCTCGCGGCTGTCTCCCTGCGACATCCAGCGCCCTTGAAGTGTGGGAAC 237
Qy 1248 TCAGGCTCTCTGGAAGAAACCTGGGCACTGGAGGCTCAG--CTGGGCGCAATTGATTGAGC 1306
Db 236 TCAGGCTCTCTGGAAGAAACCTGGGCACTGGAGGCTCAGCTGGGCGCAATTGATTGAGC 177
Qy 1307 CTTTGAAGGAGATGAGGCTGGCTCTTGTGAAGCCAGCAGAGGCTGAAGAACTCAGGCTT 1366
Db 176 CTTTGAAGGAGATGAGGCTGGCTCTTGTGAAGCCAGCAGAGGCTGAAGAACTCAGGCTT 117
Qy 1367 CCTGATATCAGAGCCGCCCTCCCATCTTCTCTGTAAGAAACAACCCCAACCCCATTC 1426
Db 116 CCTTATATCAGAGCCGCCCTCCSBRTTCTCTCTTAATTAACAAC-----C 68
Qy 1427 TACCCCCCAATGCACTTCACTCTGTGTCTCAGAGTATTAGCCTCAGACTCTTCTTT 1486
Db 67 TACCCCCCAATGCACTTCACTCTGTGTCTCAGAGTATTAGCCTTAACTCTCTTTAT 8
Qy 1487 TATTTGT 1492
Db 7 TTTTGT 2

RESULT 15
AL562353 860 bp mRNA linear EST 16-FEB-2001
LOCUS AL562353 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC007YE11 3
DEFINITION prime, mRNA Sequence.
ACCESSION AL562353
VERSION AL562353.1 GI:12910688
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 860)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
```

```
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source
1..860
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC007YE11"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 160 a 259 c 234 g 199 t 8 others
ORIGIN
Query Match 52.3%; Score 812.4; DB 9; Length 860;
Best Local Similarity 96.6%; Pred. No. 1.3e-141;
Matches 831; Conservative 8; Mismatches 20; Indels 1; Gaps 1;

Qy 489 TGCACAGACACCCCTGATGCTCTTCACTGCGCCGTAACCGGCAAGGCGCTCAGGCGCACT 548
Db 860 TGAACAGACACCCCTGATGCTCTTCACTGCGCCGTAACCGGCAAGGCGCTCAGGCGCACT 801
Qy 549 GGCAGCTCATGAAGAGATTAACCTGCTGAGAGACAGACAGTGAAGCGCGCCCAAG 608
Db 800 GGCAGCTCATGAAGAGATTAACCTGCTGAGAGACAGACAGTGAAGCGCGCCCAAG 741
Qy 609 GGGACCAAGTGTGAACCTCTCTGATGAGAGAGACCTGATTGATGACAGTAAGCTCAAG 668
Db 740 GGGACCAAGTGTGAACCTCTCTGATGAGAGAGACCTGATTGATGACAGTAAGCTCAAG 681
Qy 740 GGGACCAAGTGTGAACCTCTCTGATGAGAGAGACCTGATTGATGACAGTAAGCTCAAG 681
Db 669 ACATCGAGATGAGGCTCTGGAACATGAGCTGATGCTGACCGGCGCAAGAGCAG 728
Db 680 ACATCGAGATGAGGCTCTGGAACATGAGCTGATGCTGACCGGCGCAAGAGCAG 621
Qy 729 AGATTGCGCAGCTGGAACAGAACTGCAATAGTGCAGGTGTATATGACAGCAATCAGAG 788
Db 620 AGATTGCGCAGCTGGAACAGAACTGCAATAGTGCAGGTGTATATGACAGCAATCAGAG 561
Qy 789 GCATGAGGCTCTCGGACTTCCAGCAACAGACACTGSCAGTGTGCGGAGCGCATGTGC 848
Db 560 GCATGAGGCTCTCGGACTTCCAGCAACAGACACTGSCAGTGTGCGGAGCGCATGTGC 501
Qy 849 GGTACTGATGTCGCTCGGTGATGATCACTGCGGAGCAACCAAGAGTAACAGATTG 908
Db 500 GGTACTGATGTCGCTCGGTGATGATCACTGCGGAGCAACCAAGAGTAACAGATTG 441
Qy 908 ATGTGAGCTGTCTCTGAGAGGTCGGGCTTGAAGATATCCCGGAAACAAAGTGTCAATCA 968
Db 440 ATGTGAGCTGTCTCTGAGAGGTCGGGCTTGAAGATATCCCGGAAACAAAGTGTCAATCA 381
Qy 969 AGCTGAAGAACAGGATATTTCTTCAATGCAATGAGGATGAGCGGCCATCTCAATCG 1028
Db 380 AGCTGAAGAACAGGATATTTCTTCAATGCAATGAGGATGAGCGGCCATCTCAATCG 321
Qy 1029 ATGACGCGCGGTGCTCTGTGGCTCAAAATGGCGCTCAGCAACAACTCTGTGTGAGAGA 1088
Db 320 ATGACGCGCGGTGCTCTGTGGCTCAAAATGGCGCTCAGCAACAACTCTGTGTGAGAGA 261
Qy 1089 TGGCAGGCTGCGATTCGTTCTTATCAACAGGACTTATGCGCCCTCATCAGGCGCTG 1148
Db 260 TGGCAGGCTGCGATTCGTTCTTATCAACAGGACTTATGCGCCCTCATCAGGCGCTG 201
```


Oy	1149	AGGCTGCGCAAGATACACACCA	CAGTGAAGAAATGGTGCACAGACTCGTGGGACCTCTCCGGC	1208
Db	200	AGGCTGCGCAAGATACACCA	CAGTGAAGAAATGGTGCACAGACTCGTGGGACCTCTCCGGC	141
Oy	1209	CTGATTTCCCTCCACACTCCAG	CCCCCTTGAAGCTGGGAACTCAGGCTCTCTGGAAAAACCTG	1268
Db	140	CTGAAAAAACCTGGAACTCCAG	CCCCCTTGAAGCTGGGAACTCAGGCTCTCTGGAAAAACCTG	81
Oy	1269	GGCAGTGGGAGGCTCAGCTCG	GGGACATTGATTTGAGCTTTGAGGGAGAGATAGG-GCTG	1327
Db	80	GGCAGTGGGAGGCTCAGCTCG	GGGACATTGATTTGAGCTTTGAGGGAGAGATAGGAGCTG	21
Oy	1328	GCCTTTGTGAAGCCAGCAGA	1347	
Db	20	GCCTTTGTGAAGCCAGCAGA	1	

Search completed: March 27, 2003, 07:47:14
Job time : 2780 secs

This Page Blank (uspto)

PT treat or diagnose, e.g. inflammation and tumors, and to identify its
PT specific modulators, to regulate MEK kinase activity -
XX Claim 20; Page 64-67; 78pp; English.
XX
XX The invention provides MEK kinase (MEKK) interacting forkhead associated
CC (FHA) protein (MIF1). MIF1 is useful for screening for specific
CC modulators (potential therapeutic agents) and to reduce MEKK activity in
CC cells. Antibodies specific to MIF1 are useful as diagnostic immunoassay
CC reagents to detect expression of MIF1, for purification of MIF1 and as
CC therapeutic (ant)agonists. The MIF1 nucleic acids are useful for:
CC recombinant production of MIF1, either in cultured cells or in vivo (gene
CC therapy); as source of probes and primers for detecting or quantifying
CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source
CC of therapeutic antisense sequences (used to increase MEKK activity in
CC cells); and to identify inhibitors of MIF1. Regulation of MEKK activity
CC via MIF1 regulation is useful for treatment of inflammation, asthma,
CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic
CC syndrome, neurodegeneration, tumors, angiogenesis, rheumatoid arthritis,
CC psoriasis or persistent viral infections. The present sequence represents
CC the MIF1 protein (plasmid pcM480).
XX
XX Sequence 390 AA:
SQ
Query Match 100.0%; Score 1996; DB 21; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.2e-171;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSARGVEPGRCSGSPSSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 60
DB 1 NSARGVEPGRCSGSPSSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 60
QY 61 GWRKPRANDLLINAVLQTNLDLTSVHLGVKFSRFTLRVEQERWYALLYDPVSKLACQAM 120
DB 61 GWRKPRANDLLINAVLQTNLDLTSVHLGVKFSRFTLRVEQERWYALLYDPVSKLACQAM 120
QY 61 GWRKPRANDLLINAVLQTNLDLTSVHLGVKFSRFTLRVEQERWYALLYDPVSKLACQAM 120
DB 61 GWRKPRANDLLINAVLQTNLDLTSVHLGVKFSRFTLRVEQERWYALLYDPVSKLACQAM 120
QY 121 ROLHPEALTAIOSKALFSAEBQLSKVGSTSQPTLETFODLHHRPDAFYLAFTAKALQ 180
DB 121 ROLHPEALTAIOSKALFSAEBQLSKVGSTSQPTLETFODLHHRPDAFYLAFTAKALQ 180
QY 121 ROLHPEALTAIOSKALFSAEBQLSKVGSTSQPTLETFODLHHRPDAFYLAFTAKALQ 180
DB 121 ROLHPEALTAIOSKALFSAEBQLSKVGSTSQPTLETFODLHHRPDAFYLAFTAKALQ 180
QY 181 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKMDEVLHEHLMVADRQ 240
DB 181 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKMDEVLHEHLMVADRQ 240
QY 241 KREIRQLEBELHKQVLDVSTIGMSSPDPDNOTLAVLRGMVRYLMRSREITLGRATKDN 300
DB 241 KREIRQLEBELHKQVLDVSTIGMSSPDPDNOTLAVLRGMVRYLMRSREITLGRATKDN 300
QY 301 QIDVDLSLEGPMKISRKQGVTKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 360
DB 301 QIDVDLSLEGPMKISRKQGVTKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 360
QY 361 VEIASLRFFVLINODLIALIRAEAAKITPQ 390
DB 361 VEIASLRFFVLINODLIALIRAEAAKITPQ 390
RESULT 2
AAV77555
ID AAV77555 standard; Protein; 462 AA.
XX
XX AAV77555;
XX
XX 08-MAY-2000 (first entry)
XX
XX Human MIF1 protein (plasmid pcM577).
XX
XX MEK kinase; MEKK interacting forkhead associated protein; MIF1; MEKK;
KM FHA protein; forkhead associated protein; tumor; angiogenesis; human;
KM psoriasis; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;
KM vasotrophic; neuroprotective; antiarthritic; antiviral.
XX
XX Homo sapiens.
OS

XX
XX WO200005362-A1.
XX
XX 03-FEB-2000.
XX
XX 21-JUL-1999; 99WO-EP05142.
XX
XX 21-JUL-1998; 98US-0093590.
XX
XX (RHON) RHONE-POULENC RORER SA.
XX
XX Marcieureau C, Multon M, Polard-Houssset V;
XX
XX WPI: 2000-195102/17.
XX
XX N-PEDB; AAZ58958.
XX
XX
XX New MEK kinase interacting forkhead associated protein (MIF1) useful to
PT treat or diagnose, e.g. inflammation and tumors, and to identify its
PT specific modulators, to regulate MEK kinase activity -
XX Claim 20; Page 70-73; 78pp; English.
XX
XX
XX The invention provides MEK kinase (MEKK) interacting forkhead associated
CC (FHA) protein (MIF1). MIF1 is useful for screening for specific
CC modulators (potential therapeutic agents) and to reduce MEKK activity in
CC cells. Antibodies specific to MIF1 are useful as diagnostic immunoassay
CC reagents to detect expression of MIF1, for purification of MIF1 and as
CC therapeutic (ant)agonists. The MIF1 nucleic acids are useful for:
CC recombinant production of MIF1, either in cultured cells or in vivo (gene
CC therapy); as source of probes and primers for detecting or quantifying
CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source
CC of therapeutic antisense sequences (used to increase MEKK activity in
CC cells); and to identify inhibitors of MIF1. Regulation of MEKK activity
CC via MIF1 regulation is useful for treatment of inflammation, asthma,
CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic
CC syndrome, neurodegeneration, tumors, angiogenesis, rheumatoid arthritis,
CC psoriasis or persistent viral infections. The present sequence represents
CC the MIF1 protein (plasmid pcM577).
XX
XX Sequence 462 AA:
SQ
Query Match 98.5%; Score 1967; DB 21; Length 462;
Best Local Similarity 99.0%; Pred. No. 2.2e-168;
Matches 386; Conservative 2; Mismatches 0; Indels 2; Gaps 1;
QY 3 AFG--GVEPGRCSGSPSSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 60
DB 73 AKGAGVEPGRCSGSPSSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 132
QY 61 GWRKPRANDLLINAVLQTNLDLTSVHLGVKFSRFTLRVEQERWYALLYDPVSKLACQAM 120
DB 133 GWRKPRANDLLINAVLQTNLDLTSVHLGVKFSRFTLRVEQERWYALLYDPVSKLACQAM 192
QY 121 ROLHPEALTAIOSKALFSAEBQLSKVGSTSQPTLETFODLHHRPDAFYLAFTAKALQ 180
DB 193 ROLHPEALTAIOSKALFSAEBQLSKVGSTSQPTLETFODLHHRPDAFYLAFTAKALQ 252
QY 181 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKMDEVLHEHLMVADRQ 240
DB 253 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKMDEVLHEHLMVADRQ 312
QY 241 KREIRQLEBELHKQVLDVSTIGMSSPDPDNOTLAVLRGMVRYLMRSREITLGRATKDN 300
DB 313 KREIRQLEBELHKQVLDVSTIGMSSPDPDNOTLAVLRGMVRYLMRSREITLGRATKDN 372
QY 301 QIDVDLSLEGPMKISRKQGVTKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 360
DB 373 QIDVDLSLEGPMKISRKQGVTKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 432
QY 361 VEIASLRFFVLINODLIALIRAEAAKITPQ 390
DB 433 VEIASLRFFVLINODLIALIRAEAAKITPQ 462

RESULT 3
ID ABB57874 standard; Protein: 578 AA.
XX ABB57874;
AC ABB57874;
XX ABB57874;
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 414.
DE Drosophila melanogaster polypeptide SEQ ID NO 414.
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200171042-A2.
FN WO200171042-A2.
XX 27-SEP-2001.
PD 27-SEP-2001.
PF 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2001; 2001WO-US09231.
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
PA (PEKE) PE CORP NY.
XX (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI, 2001-656860/75.
DR N-PSDB; ABL01977.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
FT interactions -
XX Disclosure; SEQ ID NO 414; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABLJ6175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIDO
CC at ftp.wido.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 578 AA;
Query Match 51.3%; Score 1023; DB 22; Length 578;
Best Local Similarity 55.5%; Pred. No. 3e-83;
Matches 208; Conservative 60; Mismatches 97; Indels 10; Gaps 4;
OY 16 EESSSEKKKVSAPSPVPSPBPAPAGLTGRVKKSKOPL--GVTKDLGRKPANDLLILN 73
DB 204 KPPAHERSTTSRRSRHPVPAS-----KKAQRNRGFRMGVAITDGLGWKXIDIALIIT 257
OY 74 AVLQTNDLTSLVHIGVYFS CRTFLREVOERVALLDYPVISKLACQAMROLHEBAIAAIQS 133
DB 258 GIQTNDNLRIIHRYGVFSCKFTLQLQGRWVALLVEPAVSRIAIVSAIRYLHDELVESVQR 317
OY 134 KALFEKAEQULLSKGSTSQPTLETQDILLHHNPDAFYLAARKAKLOAWOMLKQYYILE 193
DB 318 KALYSVQEEDLLGTIKSSSQPKLEQFOELLDDKNASVFYCARTAKSIQNHWMLLKQYTTLP 377
OY 194 DOTVOPLPKGD--QVLNFSPAEDLIDSCKMDRDEVTLHEHLNWADRRQGRGRIROLEQH 252
DB 378 DOSVRFPIYGTDQQPLSFSAEDOIIFHDINLRPERDELNERKRLADRNRGNRIIRLENELS 437
OY 253 KWQVLVDST-TGMSSPDFDNQTLAVLRGMVRVLYMRSEKITTLGRATKDNOQDIVDSLSEGP 311

[illegible]

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0138119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145931.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149425.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.0%; Score 160; DB 21; Length 465;
Best Local Similarity 27.0%; Pred. No. 1,1e-05;
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;

Oy	211	DADDLID-DSKAK-----DMDDEVLEHLMADRRCKEIRLOEDELHKQVLVDSING	263
Db	279	DDEEEDIDAMIRKTLNVLVDDSDSCFNREEMNMSKPHRALIGLEQ-----CTRFS	329
Oy	264	MSSPFDNQTGLAVLRGRWRYLRSREITLGRATKNOIDVDLSLEGPAKISRKOGVIR	323
Db	330	MORALMFHAIALVLRHPDSDKPHFRKEVITIGRSSGLANDIDIGKYNYSKISRRAALVR	389
Oy	334	LKNNGDFFIANEGRRPIYIDGRFVLGSKMRLLNNSVETIASLRFVFLINDLI	377
Db	390	LENYGSFSLKNIKGKHILVNGGLDRGQIVTLTSCSSINIRGITFVFINKKEAV	443
RESULT 5			
AAG40260			
ID	AAG40260 standard; Protein; 465 AA.		
XX	AC	AAG40260;	
XX	DE	18-OCT-2000 (first entry)	
XX	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 49928.	
XX	KM	Protein identification; signal transduction pathway; metabolic pathway;	
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	OS	termination sequence.	
XX	OS	Arabidopsis thaliana.	
PN	EPI033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-0301439.		
XX	XX	25-FEB-1999; 990US-0121825.	
PR	05-MAR-1999;	990US-0123180.	
PR	03-MAR-1999;	990US-0123548.	
PR	23-MAR-1999;	990US-0125788.	
PR	29-MAR-1999;	990US-0126264.	
PR	25-MAR-1999;	990US-0126785.	
PR	01-APR-1999;	990US-0127462.	
PR	06-APR-1999;	990US-0128234.	
PR	08-APR-1999;	990US-0128714.	
PR	16-APR-1999;	990US-0129845.	
PR	19-APR-1999;	990US-0130077.	
PR	21-APR-1999;	990US-0130449.	
PR	23-APR-1999;	990US-0130510.	
PR	28-APR-1999;	990US-0130891.	
PR	30-APR-1999;	990US-0131449.	
PR	04-MAY-1999;	990US-0132407.	
PR	05-MAY-1999;	990US-0132484.	
PR	05-MAY-1999;	990US-0132485.	
PR	06-MAY-1999;	990US-0132486.	
PR	06-MAY-1999;	990US-0132487.	
PR	07-MAY-1999;	990US-0132863.	
PR	11-MAY-1999;	990US-0134256.	
PR	14-MAY-1999;	990US-0134218.	
PR	14-MAY-1999;	990US-0134219.	
PR	14-MAY-1999;	990US-0134221.	
PR	14-MAY-1999;	990US-0134370.	
PR	18-MAY-1999;	990US-0134368.	
PR	19-MAY-1999;	990US-0134341.	
PR	20-MAY-1999;	990US-0135424.	
PR	21-MAY-1999;	990US-0135153.	
PR	24-MAY-1999;	990US-0135629.	
PR	25-MAY-1999;	990US-0136021.	
PR	27-MAY-1999;	990US-0136392.	
PR	28-MAY-1999;	990US-0136782.	
PR	01-JUN-1999;	990US-0137222.	
PR	03-JUN-1999;	990US-0137528.	
PR	04-JUN-1999;	990US-0137502.	
PR	07-JUN-1999;	990US-0137724.	
PR	08-JUN-1999;	990US-0138094.	
PR	10-JUN-1999;	990US-0138540.	
PR	14-JUN-1999;	990US-0138847.	
PR	14-JUN-1999;	990US-0139119.	
PR	16-JUN-1999;	990US-0139452.	
PR	17-JUN-1999;	990US-0139453.	
PR	18-JUN-1999;	990US-0139492.	
PR	18-JUN-1999;	990US-0139454.	
PR	18-JUN-1999;	990US-0139455.	
PR	18-JUN-1999;	990US-0139456.	
PR	18-JUN-1999;	990US-0139457.	
PR	18-JUN-1999;	990US-0139457.	
PR	18-JUN-1999;	990US-0139458.	
PR	18-JUN-1999;	990US-0139459.	
PR	18-JUN-1999;	990US-0139460.	
PR	18-JUN-1999;	990US-0139461.	
PR	18-JUN-1999;	990US-0139462.	
PR	18-JUN-1999;	990US-0139463.	
PR	18-JUN-1999;	990US-0139750.	
PR	18-JUN-1999;	990US-0139763.	
PR	21-JUN-1999;	990US-0139817.	
PR	22-JUN-1999;	990US-0139899.	
PR	23-JUN-1999;	990US-0140353.	
PR	24-JUN-1999;	990US-0140354.	
PR	28-JUN-1999;	990US-0140823.	
PR	29-JUN-1999;	990US-0140991.	
PR	30-JUN-1999;	990US-0141287.	
PR	01-JUL-1999;	990US-0141287.	
PR	01-JUL-1999;	990US-0142154.	
PR	02-JUL-1999;	990US-0142055.	
PR	06-JUL-1999;	990US-0142390.	
PR	08-JUL-1999;	990US-0142803.	
PR	09-JUL-1999;	990US-0142920.	
PR	12-JUL-1999;	990US-0142977.	
PR	13-JUL-1999;	990US-0143542.	
PR	14-JUL-1999;	990US-0143624.	
PR	15-JUL-1999;	990US-0144005.	
PR	16-JUL-1999;	990US-0144085.	
PR	16-JUL-1999;	990US-0144086.	
PR	19-JUL-1999;	990US-0144325.	
PR	19-JUL-1999;	990US-0144331.	
PR	19-JUL-1999;	990US-0144332.	
PR	19-JUL-1999;	990US-0144333.	
PR	19-JUL-1999;	990US-0144334.	
PR	19-JUL-1999;	990US-0144335.	
PR	19-JUL-1999;	990US-0144336.	
PR	20-JUL-1999;	990US-0144352.	
PR	20-JUL-1999;	990US-0144632.	
PR	20-JUL-1999;	990US-0144884.	
PR	21-JUL-1999;	990US-0144814.	
PR	21-JUL-1999;	990US-0145086.	
PR	21-JUL-1999;	990US-0145088.	
PR	22-JUL-1999;	990US-0145085.	
PR	22-JUL-1999;	990US-0145087.	
PR	22-JUL-1999;	990US-0145088.	
PR	22-JUL-1999;	990US-0145087.	
PR	22-JUL-1999;	990US-0145087.	
PR	22-JUL-1999;	990US-0145089.	
PR	22-JUL-1999;	990US-0145192.	
PR	23-JUL-1999;	990US-0145145.	
PR	23-JUL-1999;	990US-0145218.	
PR	23-JUL-1999;	990US-0145219.	
PR	23-JUL-1999;	990US-0145224.	
PR	26-JUL-1999;	990US-0145276.	
PR	27-JUL-1999;	990US-0145913.	
PR	27-JUL-1999;	990US-0145918.	
PR	27-JUL-1999;	990US-0145919.	
PR	27-JUL-1999;	990US-0145918.	
PR	28-JUL-1999;	990US-0145951.	
PR	02-AUG-1999;	990US-0146386.	
PR	02-AUG-1999;	990US-0146388.	
PR	02-AUG-1999;	990US-0146389.	
PR	03-AUG-1999;	990US-0147038.	
PR	04-AUG-1999;	990US-0147204.	
PR	04-AUG-1999;	990US-0147302.	
PR	05-AUG-1999;	990US-0147192.	
PR	05-AUG-1999;	990US-0147260.	
PR	06-AUG-1999;	990US-0147303.	

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.0%; Score 160; DB 21; Length 465;
Best Local Similarity 27.0%; Pred. No. 1,1e-05;
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;

QY 211 DAEDID-DSKUK-----DMRDEVLEHMLWADRQRKEIRQLQEOLHKMQVLVDSTIG 263
Db 279 DGEERIDIDAMIRKLNLYPDDSDSCFNRBEKNNKXHPHALGLBQ-----CIRTS 329
QY 264 MSSPDENQTLAVLRGRWRYLMRSREITLGRATKNOIDVDLSLEGPAWKISRQGVYK 323
Db 330 MGRALMFHGAIALVHCPDSKHFVRKREVIIGRSSGGLWVIDLGKYNVGSKISRQALVK 389
QY 324 LKNGDFFPIANEGRRPYIIDGRPVLCGSKWLSNNSVVEIASLRVFLINDLI 377
Db 390 LENYSFSLKXNGKHQHIWNGKLDRGQIVTLTSCSINIRGITVEFINKKAV 443

RESULT 6
AAG07417
ID AAG07417 standard; Protein; 479 AA.
XX AAG07417;
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4563.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hydridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139753.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142920.
 PR 13-JUL-1999; 99US-0142977.
 PR 14-JUL-1999; 99US-0143542.
 PR 15-JUL-1999; 99US-0143624.
 PR 16-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145115.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149358.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159337.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 16-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query March 8.0%; Score 160; DB 21; Length 479;
 Best Local Similarity 27.0%; Pred. No. 1.1e-05;

Matches	47;	Conservative	36;	Mismatches	75;	Indels	16;	Gaps	3
Qy	211	DAEDLID-DSKUK-----DMRDEVLEHLEIMVADRQKREIRQLQEOLHKQVLYDSITG	263						
Db	293	DDEEEDIDAMLRKUNLVDPDDSDSCFNREEMNMSKPRHALIGLEQ-----CTRTS	343						
Qy	264	WSSPDFDNQTLAVLRGRWVRYLMRSREITLGATKQNDIVDLSLEGPMKISRKGVIK	323						
Db	344	MORALMFHGAIVLHCPDPSKHPVRKREVIIGRSSGGLVNDIDIGKYNYSKISRQALVK	403						
Qy	324	LKNNGGEFLANGRRPIYIDGRVULGSKMRKLSNNSVVEIASIRFVFLINODLI	377						
Db	404	LENYGSFSLKNGKHILVNGCKLDRGQIVTLTSCSSINIRGITTFVKINKEAV	457						
RESULT 7									
XX	ID	AAG40259	standard; Protein; 479 AA.						
XX	AC	AAG40259;							
XX	DT	18-OCT-2000	(first entry)						
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 49927.							
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;							
KW	hybridisation assay; genetic mapping; gene expression control; promoter;								
KW	termination sequence.								
OS	Arabidopsis thaliana.								
XX	PN	EP103405-A2.							
XX	PD	06-SEP-2000.							
XX	PF	25-FEB-2000; 2000EP-0301439.							
PR	25	FEB-1999; 99US-0121825.							
PR	05	MAR-1999; 99US-0123180.							
PR	09	MAR-1999; 99US-0123548.							
PR	23	MAR-1999; 99US-0125788.							
PR	25	MAR-1999; 99US-0126264.							
PR	29	MAR-1999; 99US-0126785.							
PR	01	APR-1999; 99US-0127462.							
PR	06	APR-1999; 99US-0128234.							
PR	08	APR-1999; 99US-0128714.							
PR	16	APR-1999; 99US-0129845.							
PR	19	APR-1999; 99US-0130077.							
PR	21	APR-1999; 99US-0130449.							
PR	23	APR-1999; 99US-0130510.							
PR	23	APR-1999; 99US-0130891.							
PR	28	APR-1999; 99US-0131449.							
PR	30	APR-1999; 99US-0132048.							
PR	30	APR-1999; 99US-0132407.							
PR	04	MAY-1999; 99US-0132484.							
PR	05	MAY-1999; 99US-0132485.							
PR	06	MAY-1999; 99US-0132486.							
PR	06	MAY-1999; 99US-0132487.							
PR	07	MAY-1999; 99US-0132863.							
PR	11	MAY-1999; 99US-0134256.							
PR	14	MAY-1999; 99US-0134218.							
PR	14	MAY-1999; 99US-0134219.							
PR	14	MAY-1999; 99US-0134221.							
PR	14	MAY-1999; 99US-0134370.							
PR	18	MAY-1999; 99US-0134768.							
PR	19	MAY-1999; 99US-0134941.							
PR	20	MAY-1999; 99US-0135124.							
PR	21	MAY-1999; 99US-0135353.							
PR	24	MAY-1999; 99US-0135629.							
PR	25	MAY-1999; 99US-0136021.							
PR	27	MAY-1999; 99US-0136392.							
PR	28	MAY-1999; 99US-0136782.							
PR	01	JUN-1999; 99US-0137222.							

PR	03-JUN-1999	99US-01375728
PR	04-JUN-1999	99US-01377502
PR	07-JUN-1999	99US-01377224
PR	08-JUN-1999	99US-01380994
PR	10-JUN-1999	99US-01385400
PR	10-JUN-1999	99US-01388477
PR	14-JUN-1999	99US-01391119
PR	16-JUN-1999	99US-01394512
PR	16-JUN-1999	99US-01394533
PR	17-JUN-1999	99US-01394992
PR	18-JUN-1999	99US-01394554
PR	18-JUN-1999	99US-01394655
PR	18-JUN-1999	99US-01394660
PR	18-JUN-1999	99US-01394661
PR	18-JUN-1999	99US-01394662
PR	18-JUN-1999	99US-01394663
PR	18-JUN-1999	99US-01397503
PR	18-JUN-1999	99US-01397633
PR	21-JUN-1999	99US-01396177
PR	22-JUN-1999	99US-01398979
PR	23-JUN-1999	99US-01403553
PR	23-JUN-1999	99US-01403554
PR	24-JUN-1999	99US-01406095
PR	28-JUN-1999	99US-01408223
PR	29-JUN-1999	99US-01409921
PR	30-JUN-1999	99US-01412877
PR	01-JUL-1999	99US-01418442
PR	01-JUL-1999	99US-01421554
PR	02-JUL-1999	99US-01422955
PR	06-JUL-1999	99US-01422950
PR	08-JUL-1999	99US-01428033
PR	09-JUL-1999	99US-01429200
PR	12-JUL-1999	99US-01429277
PR	13-JUL-1999	99US-01435244
PR	14-JUL-1999	99US-01435244
PR	15-JUL-1999	99US-01440005
PR	15-JUL-1999	99US-01440855
PR	16-JUL-1999	99US-01440866
PR	19-JUL-1999	99US-01443231
PR	19-JUL-1999	99US-01443232
PR	19-JUL-1999	99US-01443322
PR	19-JUL-1999	99US-01443324
PR	19-JUL-1999	99US-01443334
PR	19-JUL-1999	99US-01443344
PR	20-JUL-1999	99US-01443352
PR	20-JUL-1999	99US-01443525
PR	20-JUL-1999	99US-01446323
PR	20-JUL-1999	99US-01446884
PR	21-JUL-1999	99US-01448114
PR	21-JUL-1999	99US-01450866
PR	21-JUL-1999	99US-01450888
PR	22-JUL-1999	99US-01450887
PR	22-JUL-1999	99US-01450887
PR	22-JUL-1999	99US-01450889
PR	22-JUL-1999	99US-01451922
PR	23-JUL-1999	99US-01451455
PR	23-JUL-1999	99US-01452518
PR	23-JUL-1999	99US-01452524
PR	23-JUL-1999	99US-01452524
PR	26-JUL-1999	99US-01452576
PR	27-JUL-1999	99US-01452573
PR	27-JUL-1999	99US-01452578
PR	27-JUL-1999	99US-01452578
PR	28-JUL-1999	99US-01452571
PR	02-AUG-1999	99US-01463586
PR	02-AUG-1999	99US-01463586
PR	02-AUG-1999	99US-01463589
PR	03-AUG-1999	99US-01470389
PR	04-AUG-1999	99US-01472704
PR	05-AUG-1999	99US-01473792
PR	05-AUG-1999	99US-01473792

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.0%; Score 160; DB 21; Length 479;

Best Local Similarity 27.0%; Pred. No. 1,1e-05;
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;
Oy 211 DAEDLID-DSKLK-----DMRDEVLHEHLMVADRQKREIROLEQELHKKQVLVDSITG 263
Db 293 DGEERIDIDAMIRKINLVDPDSDCFNREWNMSGHPHALGLEQ-----CTRIS 343
Oy 264 MSSPDFDQUTLAVLRGRVRIAMRSREITLGATKDNQIDVDLSLEGPAWKISRQGVYK 323
Db 344 MQRALMFQALAVLHCPDSKHFVRKREYITGRSSGGLAVDIDLGKYNYSKISRQALVK 403
Oy 324 LKNGDFFIANEGRRPIYIDGRPVLCGSKMRLSNNSVVEIASLRFVLINDPLI 377
Db 404 LENVGSFSLKMLGKHQIILVNGKGLDRGQIVLTSKSSINIRGITVFVFINKKAV 457
RESULT 8
AAG07416
ID AAG07416 standard; Protein; 490 AA.
XX
AC AAG07416;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4562.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
ED 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0153363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254057.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Baraah SC, Ruben SM;
 XX
 DR WPI; 2001-502630/55.
 DR N-PSDB; AAK87783.
 XX
 PT Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases -
 XX
 PS Claim 11; SEQ ID NO 1359; 986bp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a digestive system antigen of
 CC the invention.
 CC
 XX
 XX
 SQ Sequence 108 AA;
 Query Match 6.2%; Score 123.5; DB 22; Length 108;
 Best Local Similarity 71.1%; Pred. No. 0.0027;
 Matches 27; Conservative 3; Mismatches 5; Indels 3; Gaps 2;
 QY 315 ISRKQGVTKLNKNGDFLANGRPP-IYIDGRPYLCGS 351
 : |||||
 Db 64 LSASTGVTKLNKNGDFLANGRPPSTSMGR--VCGS 99
 : |||||
 RESULT 11
 AAU19989
 ID AAU19989 standard; Protein; 108 AA.
 XX
 AC AAU19989;
 XX
 DT 06-DEC-2001 (first entry)
 XX
 DE Human liver associated polypeptide #20.
 XX

KW Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytosolic; cardiant; neuroprotective;
KW cerebroprotective; antiproliferative; neurotoxic; fungicide; cancer;
KW ophthalmological; vulnereary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX Homo sapiens.
XX OS
XX WO200155355-A1.
XX PD
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01351.
XX PE
XX 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 11-JUL-2000; 2000US-0217497.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226881.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246612.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-457728/49.
XX
XX N-PSDB; AAS31700.
XX
XX Isolated nucleic acid molecule encoding a human liver related protein
PT is used in preventing, treating or ameliorating disorders of the liver
PT particularly cancer of the liver -
XX
XX Claim 11; SEQ ID No 176; 526pp; English.
XX
XX Sequences AAU19970-AAU20115 represent the liver associated polypeptides
CC of the invention. Liver associated polypeptides and their associated
CC polynucleotides are useful in the diagnosis, treatment and prevention of
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,
CC cats, dogs, chickens or sheep. A pathological condition can be determined
CC by detecting the presence or absence of a mutation in a liver associated
CC polynucleotide. The treatable disorders include autoimmune diseases such
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as premature labour and infertility, gastrointestinal
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma and pleurisy.
CC The polypeptides can also be used to aid wound healing, to prevent skin
CC aging due to sunburn, to maintain organs before transplantation, to
CC regenerate tissues and in chemotaxis.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Query Match 6.2%; Score 123.5; DB 22; Length 108;
Best Local Similarity 71.1%; Pred. No. 0.0027;
Matches 27; Conservative 3; Mismatches 5; Indels 3; Gaps 2;
QY 315 ISRKQGVIKLKNNGDFTIANEGRRP-IYIDGRPVLCGS 351
Db 64 LSASTGVIKLKNNGDFTIANEGRRPSTSMIDGR--VCGS 99
RESULT 12
ABP40850
ID ABP40850 standard; Protein; 108 AA.
XX
XX AC ABP40850;
XX
XX 24-JUL-2002 (first entry)
XX
XX DE Human liver antigen HFLUF44, SEQ ID NO:176.
XX
XX Human liver antigen; liver disorder; hepatic disorder; infection;
KM hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
KM cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
KM autoimmune disease; Wilson's disease; primary biliary cirrhosis;
KM neoplastic disorder; cancer; tumour; portal hypertension;
KM gastrointestinal disorder; hepatitis; drug screening; gene therapy;
KM chromosome mapping; forensic analysis; antibody preparation;

KM hepatotropic; cytostatic; antiinflammatory; virucide; antibacterial;
KM fungicide; parasitic; antidote; immunosuppressive.
XX
XX OS Homo sapiens.
XX
XX US2002042096-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-0764887.
XX
XX 31-JAN-2000; 2000US-179065P.
XX
XX 04-FEB-2000; 2000US-180628P.
XX
XX 28-JUN-2000; 2000US-214886P.
XX
XX 07-JUL-2000; 2000US-216647P.
XX
XX 07-JUL-2000; 2000US-216880P.
XX
XX 11-JUL-2000; 2000US-217487P.
XX
XX 11-JUL-2000; 2000US-217496P.
XX
XX 14-JUL-2000; 2000US-218290P.
XX
XX 26-JUL-2000; 2000US-220963P.
XX
XX 26-JUL-2000; 2000US-220964P.
XX
XX 14-AUG-2000; 2000US-224518P.
XX
XX 14-AUG-2000; 2000US-224519P.
XX
XX 14-AUG-2000; 2000US-225267P.
XX
XX 14-AUG-2000; 2000US-225268P.
XX
XX 14-AUG-2000; 2000US-225270P.
XX
XX 14-AUG-2000; 2000US-225447P.
XX
XX 14-AUG-2000; 2000US-225757P.
XX
XX 14-AUG-2000; 2000US-225758P.
XX
XX 22-AUG-2000; 2000US-226868P.
XX
XX 30-AUG-2000; 2000US-228924P.
XX
XX 01-SEP-2000; 2000US-229287P.
XX
XX 01-SEP-2000; 2000US-229343P.
XX
XX 01-SEP-2000; 2000US-229344P.
XX
XX 01-SEP-2000; 2000US-229345P.
XX
XX 05-SEP-2000; 2000US-229509P.
XX
XX 05-SEP-2000; 2000US-229513P.
XX
XX 08-SEP-2000; 2000US-231413P.
XX
XX 21-SEP-2000; 2000US-234223P.
XX
XX 21-SEP-2000; 2000US-234274P.
XX
XX 25-SEP-2000; 2000US-234997P.
XX
XX 27-SEP-2000; 2000US-235834P.
XX
XX 29-SEP-2000; 2000US-236327P.
XX
XX 29-SEP-2000; 2000US-236367P.
XX
XX 29-SEP-2000; 2000US-236368P.
XX
XX 29-SEP-2000; 2000US-236369P.
XX
XX 29-SEP-2000; 2000US-236370P.
XX
XX 02-OCT-2000; 2000US-236802P.
XX
XX 02-OCT-2000; 2000US-237037P.
XX
XX 02-OCT-2000; 2000US-237038P.
XX
XX 02-OCT-2000; 2000US-237039P.
XX
XX 02-OCT-2000; 2000US-237040P.
XX
XX 13-OCT-2000; 2000US-239935P.
XX
XX 20-OCT-2000; 2000US-240960P.
XX
XX 20-OCT-2000; 2000US-241785P.
XX
XX 20-OCT-2000; 2000US-241809P.
XX
XX 01-NOV-2000; 2000US-244617P.
XX
XX 17-NOV-2000; 2000US-244299P.
XX
XX 08-DEC-2000; 2000US-251856P.
XX
XX 08-DEC-2000; 2000US-251868P.
XX
XX 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
XX
XX (ROBE/) RUBEN S M.
XX
XX (BARA/) BARASH S C.
XX
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-381944/41.
XX
XX N-PSDB; ABN90055.
XX
XX New nucleic acid encoding human liver antigens, useful for diagnosis,
PT treatment and prevention of e.g. hepatitis and hepatic cancer, also

PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001MO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li FWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL07302.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 16389; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 398 AA;
 Query Match 5.7%; Score 113.5; DB 22; Length 398;
 Best Local Similarity 29.2%; Pred. No. 0.13;
 Matches 31; Conservative 20; Mismatches 34; Indels 21; Gaps 4;
 QY 2 SARGVPERGSGSSPSSSEKKVSKASTPVPSPAPAPGLTKRVKSKQPLQTKDLG 61
 DB 102 AAGASAKFGAGSTGKPKAS-----GNAPATP-PPPPPPPKASDPVAKKPRLTKTFEY 154
 QY 62 RWKKN-----DL-----LLINAVLQTNLDTLSVHGVKFSGR 93
 DB 155 RWRKGDDQPTQTYEVDLEQCGAMVLDALIKXKNEMDPTLTFRRSCR 200
 RESULT 15
 ABB11764
 ID ABB11764 standard; peptide: 510 AA.
 XX
 AC ABB11764;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human LDL binding protein homologue, SEQ ID NO:2134.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiaesthetic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiac; vitruicide; antibacterial;
 KW antifungal; vulnery; antitumor.
 XX
 OS Homo sapiens.
 XX

PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001MO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Dymanac RT;
 DR WPI; 2001-457740/49.
 DR N-PSDB; ABA09008.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 20; Page 240-241; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 CC
 SQ Sequence 510 AA;
 Query Match 5.5%; Score 110.5; DB 22; Length 510;
 Best Local Similarity 18.3%; Pred. No. 0.36;
 Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;
 QY 1 NSARGVPERGSGSSPSSSEKKVSKASTPVPSPAPAG--LTKEVKS-- 50
 DB 102 NQGPEDGAGGFAEBDAKSKRTYVARNSEPPPTVNGEKPSKDPTEIRQSD 161
 QY 51 -----KQPLQVTKDLGRWKPANDLLINAVLQ-NDLTLSVHGVKFSRFTLEVDGRW 103
 DB 162 VGRDHRPRQKKAKGKGS-----ITLLMQTLNTLSTEEKLAAKCKYAELEBHR 215

```

Qy 104 YALYDPIVSKLACQAMRQLHPEATAIOSKALFSA----- 140
Db 216 NSQOMKLLQKKQSOLOVE--KDHIRGEHSKAVLARSKLESICRELQRHNRSLKEGVOR 273
Qy 141 --EEOLLSKVSTSQPTLETFPQDILHRHPDAFYLAFTAKALQAHWQ-----LMKOY 190
Db 274 AREEEKKEVTSHFQVTLNDIQLOMEQHNE-----RNSKLNQENMELAEKLIQOYE 328
Qy 191 LLEDQTVQPLPKGDVNLNFSDAEDLIDSKLKMDRDEVLEHELMVADRQKREIROEOE 250
Db 329 LREEHI-----DKVFKHKDLOOQLVDAKLOAOEMLKE-----AERHQRE----- 369
Qy 251 LHKMGVYVDSITGMSSPDNDOTLAVLRGMVRYLMRSREI--TLGRATKQNDIVDLSL 308
Db 370 --KDFLLKEAVESQRMCELMKQOETHLKQOALYTEKEFEFONTLSSKS-----EFTTF 422
Qy 309 EGPWKISRKQGVIKLKNNGDFPIANEGRRPIYIDGRPVLGSGKWRLSNNSVVEIASLRF 368
Db 423 KOEMEKMTKK--IKKLEKETTMY-----RSRWESSNKALLEMABEKT 462
Qy 369 V 369
Db 463 V 463

```

Search completed: March 27, 2003, 05:19:20
 Job time : 75 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:18:00 ; Search time 25 Seconds
(without alignments)
458.998 Million cell updates/sec

Title: US-09-744-125A-2
Perfect score: 1996
Sequence: 1 NSARGVPERGSGSEPPSS.....LINDLILIRAEAKITPQ 390

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110.5	5.5	530	4	US-08-979-608A-8
2	107	5.4	557	4	US-08-979-608A-5
3	102	5.1	10182	4	US-09-134-001C-3159
4	100	5.0	15281	2	US-08-471-119A-2
5	99.5	5.0	496	2	US-08-463-418-2
6	99.5	5.0	1377	2	US-08-308-818-4
7	98.5	4.9	855	2	US-08-468-558-2
8	98.5	4.9	855	2	US-08-676-444-2
9	97.5	4.9	1048	4	US-08-887-534A-85
10	97	4.9	1168	1	US-08-620-717A-9
11	97	4.9	1232	4	US-09-592-054-2
12	95.5	4.8	503	4	US-08-740-223A-18
13	95.5	4.8	503	4	US-09-202-491-4
14	95.5	4.8	503	4	US-09-202-491-10
15	95.5	4.8	503	4	US-09-709-188-18
16	92	4.6	731	2	US-08-313-185-51
17	92	4.6	731	2	US-08-458-499-15
18	92	4.6	731	3	US-09-082-614A-51
19	92	4.6	2101	1	US-08-466-390-4
20	92	4.6	2101	1	US-08-470-950-4
21	92	4.6	2101	1	US-08-467-781-4
22	92	4.6	2101	1	US-08-195-487-4
23	92	4.6	2101	2	US-08-483-924-4
24	92	4.6	2101	4	US-09-452-294-1
25	92	4.6	2101	5	PCT-US93-06160-4
26	91.5	4.6	1274	4	US-09-095-443-2
27	91	4.6	1167	1	US-08-485-568A-6

28	91	4.6	1167	2	US-08-590-554A-6	Sequence 6, Appli
29	91	4.6	1167	2	US-09-184-223-6	Sequence 6, Appli
30	90	4.5	303	1	US-08-185-432-5	Sequence 5, Appli
31	90	4.5	737	1	US-08-185-432-2	Sequence 2, Appli
32	90	4.5	737	1	US-08-185-432-4	Sequence 4, Appli
33	90	4.5	877	3	US-08-911-853-31	Sequence 31, Appli
34	90	4.5	877	4	US-09-479-409-31	Sequence 31, Appli
35	90	4.5	877	4	US-09-479-453-31	Sequence 31, Appli
36	89.5	4.5	442	1	US-08-391-615-6	Sequence 6, Appli
37	89.5	4.5	876	1	US-08-717-515-4	Sequence 4, Appli
38	89.5	4.5	1084	1	US-08-717-515-6	Sequence 6, Appli
39	89.5	4.5	1276	1	US-08-717-515-8	Sequence 8, Appli
40	89.5	4.5	1432	3	US-08-781-891-71	Sequence 71, Appli
41	89	4.5	555	4	US-09-813-872-4	Sequence 4, Appli
42	89	4.5	1137	1	US-08-369-043-2	Sequence 2, Appli
43	89	4.5	1193	4	US-09-227-725A-4	Sequence 4, Appli
44	89	4.5	2154	2	US-08-841-349-4	Sequence 4, Appli
45	88.5	4.4	495	4	US-09-351-457-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-979-608A-8
Sequence 8, Application US/08979608A
Patent No. 6355451
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979, 608A
FILING DATE: 26-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048, 547
FILING DATE: 03-JUN-1997
APPLICATION NUMBER: US 60/031, 930
FILING DATE: 27-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-002001 (Formerly 3983/59818)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-979-608A-8
Query Match 5.5%; Score 110.5; DB 4; Length 530;

Best Local Similarity 18.3%; Pred. No. 0.063;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVEBGRSGSPSSSEKKVSKAPSTVPSPAP-----PAPG--LTKRVKKS-- 50
DB 75 NCGGDEDAQGEPAPEDEAKSRVTYVANGPEPEFPVYGEKEPEKGDPTTEIRQSD 134
QY 51 -----KQPLQVTKDGRKKRANLILLINAVIQT--NDLTSVHLGVKFSQRTLRVQERW 103
DB 135 VGDHRHRRPOEKKKAGLQKE-----ITLMQTLNTLSTPEEKLAALCKKYAEILLEHR 188
QY 104 YALLVDPVSKLACQARQHLPEALAIQSKALFSA----- 140
DB 189 NSQKMKLQKQKQSLQVQE--KDLRGESKAVLARSKLESICRELQRRNRSLKEGVOR 246
QY 141 ---EQLSLKVGSTSQPTLETPODLLHRHPDAFYLTAKALQAHWQ-----LMKQYV 190
DB 247 ABEERKREKRVTSHPQVTLNDIQLOMEQHNE-----RNSKLQENNELAERLKKLIEQYE 301
QY 191 LLEDQTVQPLPKGDQVLANSDAEDLIDSKLDMDEVLEHELMVADRQKREIRQLEGE 250
DB 302 LREEH1-----DKVFKKDLQOQLVDAKLQAOEMLKE-----AEEHORE----- 342
QY 251 LHKQVLYVDSITGSSPDNDQTLAVLRGRWRYLMSREI--TLGRATKQNDIVDSL 308
DB 343 --KDFLKEAVSQRMCMLKQOETHLKOQALYTEKEFEFQNTLSKS-----EVFTTF 395
QY 309 EGPWAKISRKQGVTKLKNNGDFIANEGRRPIYIDGRPVLCGSKWELSNNSVEIASLRF 368
DB 396 KQEMEMTKK--IKKLEKETMY-----RSRWESSNKALLLEMAEEXT 435
QY 369 V 369
DB 436 V 436

RESULT 2

US-08-979-608A-5
Sequence 5, Application US/08979608A
Patent No. 6355451
GENERAL INFORMATION:

APPLICANT: Lees, Robert S.
Lees, Robert S.
Law, Simon W.
Ajijona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,608A
FILING DATE: 26-NOV-6355451-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,547
FILING DATE: 03-JUN-1997
APPLICATION NUMBER: US 60/031,930
FILING DATE: 27-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-979-608A-5

Query Match 5.4%; Score 107; DB 4; Length 557;
Best Local Similarity 18.9%; Pred. No. 0.14; Indels 114; Gaps 17;

Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17;

QY 16 EPSSEKKKVSKA-----PSTPV-----PPSPAPAGLTKRVKKS-----KQPLQVT 57
DB 105 EPEDAEKSAVYARNGEPFGTPVNGEKEKTSKAPG--TEIIRTSDEVDGRHRRPOEKK 163
QY 58 KDLGRWKPANDLILLINAVIQT--NDLTSVHLGVKFSQRTLRVQERWYALLYDPVYSKLA 116
DB 164 KAKGLQKE-----ITLMQTLNTLSTPEEKLAALCKKYAEILLEHRNSQKMKLQKQ 217
QY 117 CQAMQHLHEALAIQSKALFSA----- 150
DB 218 SOLVQE--KDLRGESKAVLARSKLESICRELQRRNRSLKEGVQARABEERKREVT 275
QY 151 TSQPTLETPODLLHRHPDAFYLTAKALQAHWQ-----LMKQYLLLEDQTVQPLPKG 203
DB 276 HFQMTLNDIQLOMEQHNE-----RNSKLQENNELAERLKKLIEQYELREH1----- 323
QY 204 DQVNFSDAEDLIDSKLDMDEVLEHELMVADRQKREIRQLEGEHLHKQVLYVDSITG 263
DB 324 DKVFKKDLQOQLVDAKLQAOEMLKE-----AEEHORE-----KDFLKEAVS 369
QY 264 MSSPDNDQTLAVLRGRWRYLMSREI--TLGRATKQNDIVDSLLEGPAWAKISRKQGV 321
DB 370 QRMCELMQOETHLKOQALYTEKEFEFQNTLSKS-----EVFTTFKQEMEMTKK--I 422
QY 322 IKLKNNGDFIANEGRRPIYIDGRPVLCGSKWELSNNSVEIASLRF 364
DB 423 KKLKLEKETMY-----RSRWESSNKALLLEMA 447

RESULT 3

US-09-134-001C-3159
Sequence 3159, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3159
LENGTH: 10182
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 5.1%; Score 102; DB 4; Length 10182;

Best Local Similarity 24.3%; Pred. No. 28;
Matches 83; Conservative 60; Mismatches 136; Indels 62; Gaps 21;

QY 16 EPSSEKKK-----VSKAPSTVPSPAPAPAGLTKR-VKSKQPLQVTKDGRWKPANDL 70

Db 3424 EPLETAKQLOANIDQKST-----DGMTQOSVQSYERKLOEAD-----KIN 3466
Qy 71 LINVLQON-DLTSVHLGVKSCFTLREVQGRWALLYD--PVL-SILAQOAMROLPE 126
Db 3467 SINVLANNPVNNAIRTN-KVETEQINNELTQAKQGLTVYDQPLINAKTALQOOSIDNOPS 3525
Qy 127 AI-----AAIOS-KALFSAEE--QLSKVGSSTQPLETFODLHRHPDAFYLTAK-A 178
Db 3526 TTGMTAETIQVNNARQKAEQVIONANKIENAGPSVQOVSEKSVQDALSELNNAKSA 3585
Qy 179 LOAHQMLKQYLLLEDQVQPL-----PKGDVINFSDAEDLIDDSKLKMDREVL-EH 231
Db 3586 LRADKQELQAV---NQLIQPTDLNNKKPASITAYNQRVQCFNSNELNSTKTITDRILKEQ 3642
Qy 232 ELWMD-RQKREIRQLEQELHKQVLYDSITGMSPPEDNQTAVLAGRMVRYL--MRS 288
Db 3643 NPSVADVNNALNKREVOCKNEARALLQON-----KEDNSALVRAREQLQAVDVQPS 3695
Qy 289 REITLGRATKDNQIDVDLSLEGPAWKISRKQGVITLKNNGD 329
Db 3696 TE-GMTQOTKD---DYNKQQAQOEISKAQVY---DNGD 3729

RESULT 4

US-08-471-119A-2
; Sequence 2, Application US/08471119A
; Patent No. 5827706

; GENERAL INFORMATION:

; APPLICANT: Leitner, Ernst

; APPLICANT: Schneider, Elisabeth

; APPLICANT: Schoegendorfer, Kurt

; APPLICANT: Weber, Gerhard

; TITLE OF INVENTION: Cyclosporin Synthetase

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5827706artis Corporation

; STREET: 59 Route 10

; CITY: East Hanover

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07936

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/471.119A

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kasenoff, Melvyn

; REGISTRATION NUMBER: 26,389

; REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201 503 8474

; TELEFAX: 201 503 8807

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15281 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Tolypocladium niveum

; STRAIN: ATCC 34921

; US-08-471-119A-2

Query Match 5.0%; Score 100; DB 2; Length 15281;

Best Local Similarly 23.2%; Pred. No. 78;
Matches 73; Conservative 47; Mismatches 118; Indels 76; Gaps 14;
Qy 14 GSESSSEKKVSKVSPRPVPPAPAPGLTKRVKSKQPLQVTKDGRMKPAN-DLLLI 72
Db 2126 GLEPSQSVAQVNNRAQV-----FPGLE--GKQVHVGTMDRLSALSPLDLYI 2174
Qy 73 NAVIQNDLTGVHGVKSCFTLREVQGRWALLYDVLVPSKLAQAMRQ-----LHP 125
Db 2175 NSVAQ-----YPSREYLAIVE---ALVRIDGVARIFPGDMRYATATKDFLVA 2220
Qy 126 EALAIQSKALFSAEQELSKVGSSTQPLETFODLHRHPDAFYLTAKALQAHQWL 185
Db 2221 RAVHTNGSKVTRSKVQGV-----ARLEIEBELLYDP-AFFTS----- 2258
Qy 186 MKQYLLLEDQVQPLPKGDQVLFNSDAEDLIDDSKLKMDREVLHELMVADRKQREIR 245
Db 2259 LKESISEEIEHEVELPKMKKNN-----ELSSRYGAVLH-TRNNQNSRSIH 2306
Qy 246 QLEQELHKM-----QVLVDSITGMSPPEDNQTAVLAGRMVRYLMSREITLGRATKD 299
Db 2307 KINAE--SWIDPASSQMDROGLARLKENKDAESIAVINITYKTIIV-ERHIAKSLADH 2363
Qy 300 NQIDVDLSLEGPAW 313
Db 2364 DDDTHSSIDGVAW 2377

RESULT 5

US-08-463-418-2
; Sequence 2, Application US/08463418
; Patent No. 5908971

; GENERAL INFORMATION:

; APPLICANT: Van Der Straeten, Dominique et al.

; TITLE OF INVENTION: CRUCIFER ACC SYNTHASE AND USES THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,418

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/962,481

; FILING DATE: 15-OCT-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 00786/161002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 496 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-463-418-2

Query Match 5.0%; Score 99.5; DB 2; Length 496;

Best Local Similarity 20.2%; Pred. No. 0.6; Matches 68; Conservative 56; Mismatches 117; Indels 95; Gaps 12;

Qy 32 PVPSPAPAPGLTK-----RVKSKQPLQVTKDLGRKWPANDLLINAVLQNTDILTSVHL 86
Db 170 PVPSSSDNFKLTVDAAEWAKKAKKVKGLITNPSNPL---GTMLDKDTLNTL-- 224
Qy 87 GYKESCRFLREVOERWYALLY---DPVISKACQAMQHLPEALIAIOSKALPSKAE 142
Db 225 -RVFVTRKNHILVDEIVATVPAGDFV-----SVAEYVNDVIDSEVNV 268
Qy 143 QLSKVGSTSCP-----TLTFPQDL---HRHPDAFYLTAKALQAHQMLKQY 190
Db 269 DLHIVYISLSKMGLPBGFRVGIYISNDVSGARCMSSFG-----VSSQTQLMAAM 322
Qy 191 LLEDQTVQPLPKGDQVLFNSDAEDLIDSKLMDRDEVLHEHLMVAD----- 237
Db 323 LSSDQFVD-----NF-----LMSSRRLGIRHKVFTTGKKADIACLTNSAGLFAM 368
Qy 238 ---RQKREIRQLEQELHMQVLVDSITGMSP-----DENDQTLAVL 277
Db 369 MDRLHLRRRNSFESIELMHTIIDVKLVSPGSSFRCTEPGWFRCFANMDDTLHVA 428
Qy 278 RGRMYRLMRSREITLGRATKDNQIDVDLSLGPAP 313
Db 429 LGRIDFVSKNKKIYKASENDQVQNSAKKLK 464

RESULT 6
US-08-308-818-4

; Sequence 4, Application US/08308818

; Patent No. 5847077

; GENERAL INFORMATION:

; APPLICANT: Green, Michael R

; APPLICANT: Reese, Joseph C

; TITLE OF INVENTION: A No. 5847077el Fungal Multisubunit Protein

; TITLE OF INVENTION: Complex Critical for Expression of Fungal Proteins

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby PC

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: US

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/308.818

; FILING DATE: 19-SEP-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, S. Peter

; REGISTRATION NUMBER: 25,351

; REFERENCE/DOCKET NUMBER: 0342/OA404

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-52707700

; TELEFAX: 212-753-6237

; TELEK: 236687

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1377 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: D. melanogaster

; IMMEDIATE SOURCE:

; CLONE: TAFII-250
US-08-308-818-4

Query Match 5.0%; Score 99.5; DB 2; Length 1377;
Best Local Similarity 21.5%; Pred. No. 2.6;

Matches 68; Conservative 51; Mismatches 100; Indels 97; Gaps 15;

Qy 5 GGVEPGRCS-----GSEPSSEKKVSKAPSPVPPSP-----PAPGLTKVKKSK 51
Db 998 GPADPTGGEFGSYRVNPKPTQKEQESQ-PKRSVGTIDBLRLPLQPAKELLRQPK 1056
Qy 52 QPLQVTKDLGRKWPANDLLINAVLQNTDILTSVHLGV-KFS--CEFTLREVOERWYALLY 108
Db 1057 VVEEIKKLKSRV---EVIDVVRTLSTEKAKAGEGMKFSNGNFSIAHQERY----- 1107
Qy 109 DPVISKACQAMQHLPEALIAIQ---SKALFSAEQLSKVG----- 149
Db 1108 ---KECQRIFDLQNVLASSEVLSTDEASSASBSBSDLKDNENMLSNKKTSTQ 1162
Qy 150 -STSQPTLETQDL-----HRHPDAFYLTAKALQAHQML----- 186
Db 1163 LSREBEELER-OELLRQLEDEHGSPSGGAGAKGKDDPGQMLATNNQRIILRTTF 1221
Qy 187 -----KOYLLDQTVQPLPKGDQVLFNSDAEDLIDSKLMDRDEVLHEHLMVADRRQ- 240
Db 1222 RNDQKQYTRVE--TVRRQPVIDAVI-----KIRTKQEQFIKQFATLDEQOK 1267
Qy 241 ---KREIRQLEQELH 253
Db 1268 EEMKREKRRIOELRR 1283

RESULT 7

US-08-468-558-2

; Sequence 2, Application US/08468558

; Patent No. 5877280

; GENERAL INFORMATION:

; APPLICANT: Wetmur, James G.

; TITLE OF INVENTION: Cloning and Expression of Thermostable

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: United States of America

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468.558

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: MSM94-05

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 855 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-468-558-2

Query Match 4.9%; Score 98.5; DB 2; Length 855;

[illegible]

```

RESULT 8
US-08-676-444-2
Sequence 2, Application US/0867644A
Patent No. 6294325
GENERAL INFORMATION:
APPLICANT: Wetmut, James G.
TITLE OF INVENTION: CLONING AND EXPRESSION OF THERMOSTABLE
FILE OF INVENTION: MUTL GENES AND PROTEINS AND USES THEREFOR
FILE REFERENCE: MSM95-02
CURRENT APPLICATION NUMBER: US/08/676,444A
CURRENT FILING DATE: 1996-07-05
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 855
TYPE: PR1
ORGANISM: Aquifex pyrophilus
US-08-676-444-2

```

[illegible]

```

Db 435 ENAEKLLEYEYKLUKKEGIGLSKIGYKNKWGYIEVTKANVYVPBPHRRROTLISNAER 494
Qy 280 RMVYIYMSREITLIGRAITKNOQIDVDI-----SLGSPAW 313
Db 495 YTTLELORLEEKILISAOTRINELIELYELREUREEVEVLEDKVGNNAITLIGEVYIQLSNAW 554
Qy 314 KISRKQGYITKLKNNQDFFIANEGSRPI 340
Db 555 LALEKGVWVPEVNEHGYELLIEEGHVP 581

```

RESULT 9
 US-08-887-534A-85
 Sequence 85, Application US/08887534A
 Patent No. 6455323
 GENERAL INFORMATION:
 APPLICANT: Holden, David W.
 TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
 NUMBER OF SEQUENCES: 106
 CORRESPONDENCE ADDRESSES:
 ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray & Borum
 STREET: 233 South Wacker Drive/6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/887,534A
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Rin-Laures, Li-Hsien
 REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 28341/33996
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: (312) 474-6600
 INFORMATION FOR SEQ ID NO: 85:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1048 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-887-534A-85

Query Match	4.9%	Score 97.5;	DB 4;	Length 1048;
Best Local Similarity	24.6%	Pred. No. 2.7;		
Matches	45;	Conservative	35;	Mismatches 84;
				Indels 19;
				Gaps 7;
Qy	116	ACQARQLPHEALIAIGKALF	SKAEQQLSKVSGTSPLET	FODLLHR-HPDAFYUAR 174
Db	516	AVEAVALQALP---	GVNQRLLALENVKVLGEGATL	RGQDALTKQQLORDENQSLRQ 572
Qy	175	TAKALQAHQMLKQYLLLEDQ	TVQRLPRKQDVLNFSDAEDL	IDSKLMDREVLNHEIM 234
Db	573	DEQALTOQMAVTSLSNI---	TLRPL--DRIDPLDAQ----	DENKQLRLLSQRHEIQ 622
Qy	235	VADRRQKEINQLBEL-HKMQV	LVDSITG-----MSSPDPNQ	TLAULRGKRVKILMSR 289
Db	623	GQIAAHNQOIIQYQOQIL	QRQDLTLTLTGALTLPRQDEE	SSWLATRQQLQEAQSWOQRON 682
Qy	290	ETT 292		
Db	683	ETT 685		
RESULT	10			


```
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/022/999
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 333
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: TIE ligand-4
LOCATION: 1...503
OTHER INFORMATION:
US-08-740-223A-18

Query Match      4.8%; Score 95.5; DB 4; Length 503;
Best Local Similarity 21.8%; Pred. No. 1.4;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

QY 30 STVPSPAPAPGLTKRVKKS-KOPLQVTKDLGRWKPANDLLINAVLQTN--LTSVHL 86
D 52 SEPCPGPEVSRDNTLQRESLANPLHIGK-----LPTQVQKLEQALQNTQWLKLER 106
QY 87 GVKSRCFTLEVEQERWALLYDPVISKLACQAMQLHPAIAIQSALSKAEQILLS 146
D 107 AIKTLRSKLEVOQQAQONQTAPEWL-ELGTSILNQ-----TTAQIRK---LTDMEAQLLN 158
QY 147 KVGSTQPTLETF-----ODLHRHPDAFYARTAKALQAHQOLM--KOYILLEQDTV 197
D 159 QTSRMDAQMPETFTSTNKLQENOLLQROKLOLOGQNSALEKRLQALETKQ---QEELA 214
QY 198 QPLPKGDQVINFSDAEDLIDSKLDMRDEVLEHELMVADRQRKEIRQLQELHKQVL 257
D 215 SILSKAKATLNTLSRQSAALTNIERGLRG-VRHNSLSLQD--QQHSRLQ---LVLLRHL 268
QY 258 VDSITGSSPPFDNOTLAVLGRWV---RYLMRS-----REITLGATKQNOIDVDL 306
D 269 VOERANASAPAF-----IMAGEQVFQDCAEIQRSQASASGVYTIQVSNATKPRKVFCDL 322
QY 307 SLEGPAMKI--SRKQGVYKLNKNN-----GDPFIANE-----GRPIY 341
D 323 QSSGGRMTLIGRENGTVNFORWKDYQKGRGDPAGEHMLGNEVVHQLTRRAY 376

RESULT 13
US-09-202-491-4
Sequence 4, Application US/09202491
Patent No. 6432667
GENERAL INFORMATION:
APPLICANT: Valenzuela et al.
TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
FILE REFERENCE: REG330-K
CURRENT APPLICATION NUMBER: US/09/202,491
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: PCT/US97/10728
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: 60/022,999
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: 60/021,087
EARLIER FILING DATE: 1996-07-02
EARLIER APPLICATION NUMBER: 08/665,926
EARLIER FILING DATE: 1996-06-19
NUMBER OF SEQ ID NOS: 14
```

```
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 503
TYPE: PRT
ORGANISM: Homo sapiens
US-09-202-491-4

Query Match      4.8%; Score 95.5; DB 4; Length 503;
Best Local Similarity 21.8%; Pred. No. 1.4;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

QY 30 STVPSPAPAPGLTKRVKKS-KOPLQVTKDLGRWKPANDLLINAVLQTN--LTSVHL 86
D 52 SEPCPGPEVSRDNTLQRESLANPLHIGK-----LPTQVQKLEQALQNTQWLKLER 106
QY 87 GVKSRCFTLEVEQERWALLYDPVISKLACQAMQLHPAIAIQSALSKAEQILLS 146
D 107 AIKTLRSKLEVOQQAQONQTAPEWL-ELGTSILNQ-----TTAQIRK---LTDMEAQLLN 158
QY 147 KVGSTQPTLETF-----ODLHRHPDAFYARTAKALQAHQOLM--KOYILLEQDTV 197
D 159 QTSRMDAQMPETFTSTNKLQENOLLQROKLOLOGQNSALEKRLQALETKQ---QEELA 214
QY 198 QPLPKGDQVINFSDAEDLIDSKLDMRDEVLEHELMVADRQRKEIRQLQELHKQVL 257
D 215 SILSKAKATLNTLSRQSAALTNIERGLRG-VRHNSLSLQD--QQHSRLQ---LVLLRHL 268
QY 258 VDSITGSSPPFDNOTLAVLGRWV---RYLMRS-----REITLGATKQNOIDVDL 306
D 269 VOERANASAPAF-----IMAGEQVFQDCAEIQRSQASASGVYTIQVSNATKPRKVFCDL 322
QY 307 SLEGPAMKI--SRKQGVYKLNKNN-----GDPFIANE-----GRPIY 341
D 323 QSSGGRMTLIGRENGTVNFORWKDYQKGRGDPAGEHMLGNEVVHQLTRRAY 376

RESULT 14
US-09-202-491-10
Sequence 10, Application US/09202491
Patent No. 6432667
GENERAL INFORMATION:
APPLICANT: Valenzuela et al.
TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
FILE REFERENCE: REG330-K
CURRENT APPLICATION NUMBER: US/09/202,491
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: PCT/US97/10728
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: 60/022,999
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: 60/021,087
EARLIER FILING DATE: 1996-07-02
EARLIER APPLICATION NUMBER: 08/665,926
EARLIER FILING DATE: 1996-06-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 10
LENGTH: 503
TYPE: PRT
ORGANISM: Homo sapiens
US-09-202-491-10

Query Match      4.8%; Score 95.5; DB 4; Length 503;
Best Local Similarity 21.8%; Pred. No. 1.4;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

QY 30 STVPSPAPAPGLTKRVKKS-KOPLQVTKDLGRWKPANDLLINAVLQTN--LTSVHL 86
D 52 SEPCPGPEVSRDNTLQRESLANPLHIGK-----LPTQVQKLEQALQNTQWLKLER 106
QY 87 GVKSRCFTLEVEQERWALLYDPVISKLACQAMQLHPAIAIQSALSKAEQILLS 146
D 107 AIKTLRSKLEVOQQAQONQTAPEWL-ELGTSILNQ-----TTAQIRK---LTDMEAQLLN 158
```

```
QY 147 KVGSTSQPTLETF-----QDLHRHPDAFYIARTAKALQAHWQLM--KQYVLEEDQTV 197
      :      |||      :      :      :      :      :      :      :
DB 159 QTSRMDAQWPEFELSTNKLLENQLLQROKLOQOGNSALEKRLQALETKQ---QEELA 214
      :      :      :      :      :      :      :      :      :
QY 198 QLPKGDQVLNFSDAEDLIDDSKDKMDEVELEHMLVADRQKREIRQLEQELHKQVYL 257
      :      :      :      :      :      :      :      :      :
DB 215 SLISKAKALNLTLSROSALTNIERGLRG-VRHNSSLQD--QGHSLRQ---LVILRHL 268
      :      :      :      :      :      :      :      :      :
QY 258 VDSITGMSSPDFDNQTLAVLRGMV---RYLMRS-----REITLGRATKDNQIDVDL 306
      :      :      :      :      :      :      :      :      :
DB 269 VQERANASAPAF-----IMAGEQVFQDCAEIQRSASASGVYTIQVSNATKPRKVFCDL 322
      :      :      :      :      :      :      :      :      :
QY 307 SLEGPAMKI--SRKQGVILKKN-----GDFFIANE-----GRPIY 341
      :      :      :      :      :      :      :      :      :
DB 323 QSSGGRWTLIQRRNGTVNFORWMDYKQGFDPAGEHMLGNEVVAHQLTFRRAAY 376
      :      :      :      :      :      :      :      :      :
```

RESULT 15

```
US-09-709-188-18
; Sequence 18, Application US/09709188
; Patent No. 6441137
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed ligand - Vascular Intercellular Signaling Molecule
; FILE REFERENCE: REG 333-Z
; CURRENT APPLICATION NUMBER: US/09/709,188
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-709-188-18
```

Query Match 4.8%; Score 95.5; DB 4; Length 503;

Best Local Similarity 21.8%; Pred. No. 1.4;

Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

```
QY 30 STVPSPAPAPGLTRKVKKS--KQPLQVTKDILGRMKRPANDLLINAVLOTND--LTSVHL 86
      :      :      :      :      :      :      :      :      :
DB 52 SEBCPPGPEVSRDSNTLQRESLANPLHLGK---LPTQGVKQLEQALQNTQWMLKKLR 106
      :      :      :      :      :      :      :      :      :
QY 87 GYKFSQFTLREYQERWYALLYPVLSKLAQAMRQLHPEAIAIOSKALFSKAEBOQLS 146
      :      :      :      :      :      :      :      :      :
DB 107 AKITIRSKLEQVQOQMAQNOTAPML-ELGTSILNQ---TTAQIRK---LTDMEAQLLN 158
      :      :      :      :      :      :      :      :      :
QY 147 KVGSTSQPTLETF-----QDLHRHPDAFYIARTAKALQAHWQLM--KQYVLEEDQTV 197
      :      :      :      :      :      :      :      :      :
DB 159 QTSRMDAQWPEFELSTNKLLENQLLQROKLOQOGNSALEKRLQALETKQ---QEELA 214
      :      :      :      :      :      :      :      :      :
QY 198 QLPKGDQVLNFSDAEDLIDDSKDKMDEVELEHMLVADRQKREIRQLEQELHKQVYL 257
      :      :      :      :      :      :      :      :      :
DB 215 SLISKAKALNLTLSROSALTNIERGLRG-VRHNSSLQD--QGHSLRQ---LVILRHL 268
      :      :      :      :      :      :      :      :      :
QY 258 VDSITGMSSPDFDNQTLAVLRGMV---RYLMRS-----REITLGRATKDNQIDVDL 306
      :      :      :      :      :      :      :      :      :
DB 269 VQERANASAPAF-----IMAGEQVFQDCAEIQRSASASGVYTIQVSNATKPRKVFCDL 322
      :      :      :      :      :      :      :      :      :
QY 307 SLEGPAMKI--SRKQGVILKKN-----GDFFIANE-----GRPIY 341
      :      :      :      :      :      :      :      :      :
DB 323 QSSGGRWTLIQRRNGTVNFORWMDYKQGFDPAGEHMLGNEVVAHQLTFRRAAY 376
      :      :      :      :      :      :      :      :      :
```

Search completed: March 27, 2003, 05:23:22

Job time : 40 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:21:49 ; Search time 30 Seconds

(without alignments)
763.408 Million cell updates/sec

Title: US-09-744-125a-2

Perfect score: 1996
Sequence: 1 NSARGVEPRGSGSEPS...LINDLIATIRAKITPQ 390Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*\n2: /cgn2_6/ptodata/1/pubppaa/PCr_NEW_PUB pep:*\n3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*\n4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*\n5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*\n6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*\n7: /cgn2_6/ptodata/1/pubppaa/PCrUS_PUBCOMB pep:*\n8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*\n9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*\n10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*\n11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*\n12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*\n13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*\n14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123.5	6.2	108	US-09-764-887-176	Sequence 176, App
2	110.5	5.5	530	US-09-976-740-8	Sequence 8, Appli
3	110.5	5.5	530	US-09-962-055-8	Sequence 8, Appli
4	110.5	5.5	530	US-10-023-529-8	Sequence 8, Appli
5	110.5	5.5	530	US-10-023-523-8	Sequence 8, Appli
6	110.5	5.5	546	US-09-976-740-44	Sequence 44, Appli
7	110.5	5.5	546	US-10-023-529-44	Sequence 44, Appli
8	110.5	5.5	546	US-10-023-523-44	Sequence 44, Appli
9	107	5.4	557	US-09-976-740-5	Sequence 5, Appli
10	107	5.4	557	US-09-962-055-5	Sequence 5, Appli
11	107	5.4	557	US-10-023-523-5	Sequence 5, Appli
12	107	5.4	557	US-10-023-529-5	Sequence 5, Appli
13	106	5.3	519	US-09-833-790-435	Sequence 435, App
14	102.5	5.1	751	US-09-864-761-38419	Sequence 38419, A
15	100.5	5.0	475	US-09-864-761-11259	Sequence 11259, A
16	99	5.0	327	US-09-777-745-4	Sequence 4, Appli
17	97.5	4.9	1048	US-09-741-669-409	Sequence 409, App
18	97.5	4.9	1048	US-09-815-242-10062	Sequence 10062, A
19	96	4.8	558	US-09-777-745-6	Sequence 6, Appli

20	94.5	4.7	465	US-10-102-806-676	Sequence 676, App
21	94.5	4.7	1033	US-09-820-843A-75	Sequence 75, Appli
22	93.5	4.7	277	US-09-925-300-1467	Sequence 1467, Ap
23	93	4.7	302	US-10-043-487-248	Sequence 248, App
24	92.5	4.6	1312	US-09-737-178-120	Sequence 120, App
25	92	4.6	712	US-09-925-302-678	Sequence 678, App
26	91.5	4.6	526	US-09-777-745-2	Sequence 2, Appli
27	91.5	4.6	1274	US-10-020-215-2	Sequence 327, App
28	90.5	4.5	573	US-10-043-487-327	Sequence 3, Appli
29	90	4.5	1205	US-10-173-123-3	Sequence 4, Appli
30	90	4.5	1207	US-10-173-123-4	Sequence 4, Appli
31	89.5	4.5	544	US-10-174-590-274	Sequence 274, App
32	89.5	4.5	544	US-10-176-758-274	Sequence 274, App
33	89.5	4.5	544	US-10-175-737-274	Sequence 274, App
34	89.5	4.5	544	US-10-173-706-274	Sequence 274, App
35	89.5	4.5	544	US-10-175-738-274	Sequence 274, App
36	89.5	4.5	544	US-10-175-752-274	Sequence 274, App
37	89.5	4.5	544	US-10-176-482-274	Sequence 274, App
38	89.5	4.5	544	US-10-176-757-274	Sequence 274, App
39	89.5	4.5	544	US-10-176-913-274	Sequence 274, App
40	89.5	4.5	544	US-10-180-552-274	Sequence 274, App
41	89.5	4.5	544	US-10-180-557-274	Sequence 274, App
42	89.5	4.5	544	US-10-173-700-274	Sequence 274, App
43	89.5	4.5	544	US-10-174-572-274	Sequence 274, App
44	89.5	4.5	544	US-10-174-579-274	Sequence 274, App
45	89.5	4.5	544	US-10-174-582-274	Sequence 274, App

ALIGNMENTS

```
RESULT 1\nUS-09-764-887-176\nSequence 176, Application US/09764887\nPatent No. US20020042096A1\nGENERAL INFORMATION:\nAPPLICANT: Rosen et al.\nTITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies\nFILE REFERENCE: PA113\nCURRENT APPLICATION NUMBER: US/09/764,887\nCURRENT FILING DATE: 2001-01-17\nPrior application data removed - consult PALM or file wrapper\nNUMBER OF SEQ ID NOS: 658\nSOFTWARE: PatentIn Ver. 2.0\nSEQ ID NO 176\nLENGTH: 108\nTYPE: PRT\nORGANISM: Homo sapiens\nFEATURE:\nNAME/KEY: SITE\nLOCATION: (104)\nOTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids\nNAME/KEY: SITE\nLOCATION: (105)\nOTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids\nUS-09-764-887-176\nQuery Match\nBeat Local Similarly 71.1%; Pred. No. 0.0015;\nMatches 27; Conservative 3; Mismatches 5; Indels 3; Gaps 2;\nCy 315 ISRKGQVTKLKNNGDFIANGRRP-IYIDGRPVLCGS 351\nDb 64 LSASTGVTKLKNNGDFIANGRRPSTMDGR--VCGS 99\nRESULT 2\nUS-09-976-740-8\nSequence 8, Application US/09976740\nPublication No. US20020194633A1\nGENERAL INFORMATION:\nAPPLICANT: Lees, Ann M.\nAPPLICANT: Lees, Robert S.
```

```

; APPLICANT: Law, Simon W.
; APPLICANT: Atjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-740-8

Query Match          5.5% Score 110.5; DB 9; Length 530;
Best Local Similarity 18.3%; Pred. No. 0.2; Mismatches 150; Indels 111; Gaps 16;
Matches 77; Conservative 83;

QY 1 NSARGVGPGRCSGSPSSSEKKVSKAPSTVPSPSPAPG-LTKRVKKS-- 50
DB 75 NCGGPGDGAQGPAPBPAEDAKSRITYVANNGEPPEPVVYGEKPSKGPNTFEIRQSD 134
QY 51 -----KQPLQYTKDLGRKKRPAANDLLINAVIQT-NDLTSVHLGVKFSGRFTLRVQERW 103
DB 135 VGDPRHRRPOEKKKAKGLQKE-----ITLMQTLNTLSTPEKKLAALCKYAEILLEHR 188
QY 104 VALLVDPVSKLACQAMRQLHPEALIAIOSKALFSA----- 140
DB 189 NSQOKMLQKQKQSOVLQVE--KDHLSGHSKAVLARSKLESICRELQRNRSIKKEGVOR 246
QY 141 ---EEOILSKVGSSTQPLTFQDILHRRHPDAFYLAATAKALQAHWQ-----LMKQY 190
DB 247 AAEBEKREKREVSHPQVTLNDIQLQMEQHNE-----RNSKLRQENNELAERLKKLIEQYE 301
QY 191 LLEDQTVQPLPGKDQVLYNSDAEDLIDSKLKDMDDEVLEHMLVADRQKEIRQLEGE 250
DB 302 LREEH1-----DKVFKHKDLOOQLVDAKLOQAQEMLKE-----AEERHORE----- 342
QY 251 LHKQVLYVDSITGMSPPDNDQTLAVLRGMVRYLMRSREI--TLGRATKDNQIDVDLSL 308
DB 343 --KDFLKEAVESQRMCELMKQOETHLKOQALYTEKEPFPQNTLSKS-----EVFTTF 395
QY 309 EGPAMKISRKQGVILKKNNGDFFIANEGRPIYIDGRPYLCSKXWRLSNNSVVEIASLRF 368
DB 396 KQEMEMTKK--IKKLEKETTMY-----RSRWESSKALLENVAEEXT 435
QY 369 V 369
DB 436 V 436

```

```

; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
; US-09-962-055-8

Query Match          5.5% Score 110.5; DB 10; Length 530;
Best Local Similarity 18.3%; Pred. No. 0.2; Mismatches 150; Indels 111; Gaps 16;
Matches 77; Conservative 83;

QY 1 NSARGVGPGRCSGSPSSSEKKVSKAPSTVPSPSPAPG-LTKRVKKS-- 50
DB 75 NCGGPGDGAQGPAPBPAEDAKSRITYVANNGEPPEPVVYGEKPSKGPNTFEIRQSD 134
QY 51 -----KQPLQYTKDLGRKKRPAANDLLINAVIQT-NDLTSVHLGVKFSGRFTLRVQERW 103
DB 135 VGDPRHRRPOEKKKAKGLQKE-----ITLMQTLNTLSTPEKKLAALCKYAEILLEHR 188
QY 104 VALLVDPVSKLACQAMRQLHPEALIAIOSKALFSA----- 140
DB 189 NSQOKMLQKQKQSOVLQVE--KDHLSGHSKAVLARSKLESICRELQRNRSIKKEGVOR 246
QY 141 ---EEOILSKVGSSTQPLTFQDILHRRHPDAFYLAATAKALQAHWQ-----LMKQY 190
DB 247 AAEBEKREKREVSHPQVTLNDIQLQMEQHNE-----RNSKLRQENNELAERLKKLIEQYE 301
QY 191 LLEDQTVQPLPGKDQVLYNSDAEDLIDSKLKDMDDEVLEHMLVADRQKEIRQLEGE 250
DB 302 LREEH1-----DKVFKHKDLOOQLVDAKLOQAQEMLKE-----AEERHORE----- 342
QY 251 LHKQVLYVDSITGMSPPDNDQTLAVLRGMVRYLMRSREI--TLGRATKDNQIDVDLSL 308
DB 343 --KDFLKEAVESQRMCELMKQOETHLKOQALYTEKEPFPQNTLSKS-----EVFTTF 395
QY 309 EGPAMKISRKQGVILKKNNGDFFIANEGRPIYIDGRPYLCSKXWRLSNNSVVEIASLRF 368
DB 396 KQEMEMTKK--IKKLEKETTMY-----RSRWESSKALLENVAEEXT 435
QY 369 V 369
DB 436 V 436

```

```

; RESULT 3
; US-09-962-055-8
; Sequence 8, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Atjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS

```

RESULT 4
US-10-023-529-8
; Sequence 8, Application US/10023529
; Patent No. US2002012938A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-8

Query Match 5.5%; Score 110.5; DB 12; Length 530;
Best Local Similarity 18.3%; Pred. No. 0.2;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;
QY 1 NSARGVPERGCSGSEPSSEKKVSKAPSTVPSPA-----PAPG-LTKRVKKS-- 50
DB 75 NQGGEGDGAQGEPAEPDAEKSRITYVANGPEPTPVYGEKPSKGDPTTEIRQSD 134
QY 51 -----KQPLQVTKDGLGKWKRPANDLLINAVLQ-NDLTSVHLGVKFSRCFTLREVOERW 103
DB 135 VGDRHRRPQEKKAQKAGKGE-----ITLLMQTLNTLSTPEKLAALCKYAELEBHR 188
QY 104 YALLYDPVISTLACQAMQHLPEALIAIQAQSKALFSA----- 140
DB 189 NSQKQMKLQKKQSQVLQVE--KDHLRGSHSKAVLARSKLESCELOHRNRSLSKEGVQR 246
QY 141 ---BEQLISKVGSSTQPLTETFODLHHRHDPAFYLARTAKALQAHQ-----LMKQYY 190
DB 247 AREBEKKEKVTSHFOVTLNDIQLOMEQHNE-----RNSKLRQEMMELAEKLEKLEIOYE 301
QY 191 LLEDQVQPLPKGDQVLFNSDAEDLIDSKLDMDEVLEHLMVADRQREIRQLEGE 250
DB 302 LREEH-----DKVFGKDLQOOLVDAKLQQAQEMLKE-----AEERHORE----- 342
QY 251 LHKQVULVDSITGSSPPFDNQTLAVLRGWRVYLMRSREI--TLGRATKXNQIDVDLSL 308
DB 343 --KQFLKEAVESQRMCELMKQOETHLKQQLALYTEKEFEQNTLSKSS-----EVFTTF 395
QY 309 EGPAMKISKQGVILKKNQDGFPIANEGRRPIYIDGRPVLCSSKRLSNNSVVEIASLRF 368
DB 396 KQEMKMTKK--IKKLEKETMY-----RSRWSSSNKALLEVAEKEKT 435
QY 369 V 369
DB 436 V 436

; Sequence 8, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-8

Query Match 5.5%; Score 110.5; DB 12; Length 530;
Best Local Similarity 18.3%; Pred. No. 0.2;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;
QY 1 NSARGVPERGCSGSEPSSEKKVSKAPSTVPSPA-----PAPG-LTKRVKKS-- 50
DB 75 NQGGEGDGAQGEPAEPDAEKSRITYVANGPEPTPVYGEKPSKGDPTTEIRQSD 134
QY 51 -----KQPLQVTKDGLGKWKRPANDLLINAVLQ-NDLTSVHLGVKFSRCFTLREVOERW 103
DB 135 VGDRHRRPQEKKAQKAGKGE-----ITLLMQTLNTLSTPEKLAALCKYAELEBHR 188
QY 104 YALLYDPVISTLACQAMQHLPEALIAIQAQSKALFSA----- 140
DB 189 NSQKQMKLQKKQSQVLQVE--KDHLRGSHSKAVLARSKLESCELOHRNRSLSKEGVQR 246
QY 141 ---BEQLISKVGSSTQPLTETFODLHHRHDPAFYLARTAKALQAHQ-----LMKQYY 190
DB 247 AREBEKKEKVTSHFOVTLNDIQLOMEQHNE-----RNSKLRQEMMELAEKLEKLEIOYE 301
QY 191 LLEDQVQPLPKGDQVLFNSDAEDLIDSKLDMDEVLEHLMVADRQREIRQLEGE 250
DB 302 LREEH-----DKVFGKDLQOOLVDAKLQQAQEMLKE-----AEERHORE----- 342
QY 251 LHKQVULVDSITGSSPPFDNQTLAVLRGWRVYLMRSREI--TLGRATKXNQIDVDLSL 308
DB 343 --KQFLKEAVESQRMCELMKQOETHLKQQLALYTEKEFEQNTLSKSS-----EVFTTF 395
QY 309 EGPAMKISKQGVILKKNQDGFPIANEGRRPIYIDGRPVLCSSKRLSNNSVVEIASLRF 368
DB 396 KQEMKMTKK--IKKLEKETMY-----RSRWSSSNKALLEVAEKEKT 435
QY 369 V 369
DB 436 V 436

RESULT 5
US-10-023-523-8

RESULT 6
US-09-976-740-44
; Sequence 44, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:

```

; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-744-125a-2.rapb

Query Match          5.5%; Score 110.5; DB 9; Length 546;
Best Local Similarity 18.3%; Pred. No. 0.21;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVEBERGCSGSEPSSEKKVSKAPSTVPSPSPAPG-LTKRVKYS-- 50
DB 91 NCGGPEDEGAQGEPAEPEDAEKSRITYVANRGEPEPTPVNNGEKEPKSGDPNTEIRQSD 150
QY 51 -----KQPLQVTKDLGRWKRPANDLLINAVLOT-NDLTSVHLGVKFSRFTLREVOERW 103
DB 151 VQDRDHRPRQEKKKKGLGKE-----ITLMQTLNTLSTPEKLAALCKYAELEHER 204
QY 104 VALLYDPVTSKLCAQAMROLHPEAIAIOSKALFSKA----- 140
DB 205 NSQKQMKLQKKQSOQVQGE--KDHLRGSHSKAVLARSKLESICRELQHRNRSLEKEGVOR 262
QY 141 ---EEQLSKVGSISQPTLETFODLLHRHPDAFYLAATAKALQAHQ-----LMQYY 190
DB 263 AEEBEKREKREVTSHFOVTLNDIQLOMEQHNE-----RNSKLRQENNELAERLKLIQYE 317
QY 191 LLEDQTVQPLPRGQDVLPNSDAEDLIDSKLKDMDDEVLEHEIMVADRQKEIRQLEGE 250
DB 318 LREHHI-----DKVFYHKDLQOQLVDKLAQQAQEMLKE-----AEEHHQRE----- 358
QY 251 LHKQVULVDSITGMSSPDFDQTLAVLRGRMVRYLMRSREI--TLGRATKQNDIVDLSL 308
DB 359 --KDFLKEAVESQRMCELMQOQETHLKOQALALYTEKEFEFQNTLSKSS-----EVFTTF 411
QY 309 EGPAMKISRKQGVILKNNGDFFIANRGRPIYIDGRPLVCSKRLSNVVEIASLRF 368
DB 412 KQEMEMTKK--IKLEKETTWY-----RSRWSSNKALLENMAEERT 451
QY 369 V 369
DB 452 V 452

RESULT 7
US-10-023-529-44
; Sequence 44, Application US/10023529
; Patent No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
```

```

; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-529-44

Query Match          5.5%; Score 110.5; DB 12; Length 546;
Best Local Similarity 18.3%; Pred. No. 0.21;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVEBERGCSGSEPSSEKKVSKAPSTVPSPSPAPG-LTKRVKYS-- 50
DB 91 NCGGPEDEGAQGEPAEPEDAEKSRITYVANRGEPEPTPVNNGEKEPKSGDPNTEIRQSD 150
QY 51 -----KQPLQVTKDLGRWKRPANDLLINAVLOT-NDLTSVHLGVKFSRFTLREVOERW 103
DB 151 VQDRDHRPRQEKKKKGLGKE-----ITLMQTLNTLSTPEKLAALCKYAELEHER 204
QY 104 VALLYDPVTSKLCAQAMROLHPEAIAIOSKALFSKA----- 140
DB 205 NSQKQMKLQKKQSOQVQGE--KDHLRGSHSKAVLARSKLESICRELQHRNRSLEKEGVOR 262
QY 141 ---EEQLSKVGSISQPTLETFODLLHRHPDAFYLAATAKALQAHQ-----LMQYY 190
DB 263 AEEBEKREKREVTSHFOVTLNDIQLOMEQHNE-----RNSKLRQENNELAERLKLIQYE 317
QY 191 LLEDQTVQPLPRGQDVLPNSDAEDLIDSKLKDMDDEVLEHEIMVADRQKEIRQLEGE 250
DB 318 LREHHI-----DKVFYHKDLQOQLVDKLAQQAQEMLKE-----AEEHHQRE----- 358
QY 251 LHKQVULVDSITGMSSPDFDQTLAVLRGRMVRYLMRSREI--TLGRATKQNDIVDLSL 308
DB 359 --KDFLKEAVESQRMCELMQOQETHLKOQALALYTEKEFEFQNTLSKSS-----EVFTTF 411
QY 309 EGPAMKISRKQGVILKNNGDFFIANRGRPIYIDGRPLVCSKRLSNVVEIASLRF 368
DB 412 KQEMEMTKK--IKLEKETTWY-----RSRWSSNKALLENMAEERT 451
QY 369 V 369
DB 452 V 452

RESULT 8
US-10-023-523-44
; Sequence 44, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
```


;; CURRENT APPLICATION NUMBER: US/10/023,523
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 44
;; LENGTH: 546
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-023-523-44

Query Match 5.4%; Score 110.5; DB 12; Length 546;
Best Local Similarity 18.3%; Pred. No. 0.21; Indels 11; Gaps 16;
Matches 77; Conservative 83; Mismatches 150;

QY 1 NSARGVPEPGSCGSEPSSEKSKVSKAPSTVPSPSPAPG-LTKVYKKS-- 50
DB 91 NQGRBGEAGQGEPAEPDAEKSRITYVARNGEPPTPVNNGEKESKQDPTEIRQSD 150
QY 51 -----KQPLQVTKDGLRWKRPANDLLINAVLQT-NDLTSVHLGVKFSCTFTLREVOERW 103
DB 151 VQDRDHRPRQEKKAKGLGKE-----ITLMQTLNTLSTPEEKALCKKYAELLEHR 204
QY 104 YALYDPISTKLAQAMQHPAIAAQKALFESKA----- 140
DB 205 NSQKMKLQKQSQVQE--KDLRGESKAVNLSKLSLCRELQHNHNSLKEGVQR 262
QY 141 ---EEQLSKYGSTQPTLETFTQDLHHRPDAFYLAFTAKALQAHWQ-----LMKOYU 190
DB 263 AREEEKKEVTSHTQVTLNDIQLOMEQHNE-----RNSKLRQENMELAEKLIQYLEEYE 317
QY 191 LLEDDQVQPLPKGDQVILNPSDAEDLIDSKLDMDEVLHEHLMVADRROKREIRQLEOE 250
DB 318 LREHNI-----DKVFYGHKDLQQLVDAKLQQAQEMLKE-----AEERHQRE----- 358
QY 251 LHKQVLDVSTGSSPPFDNQTLAVLRGMRYLMRSREI--TLGRATKDNQIDVDLSL 308
DB 359 --KQFLLEAVEVSQRMCELMKQOETHLKQOLALYTEKEFEFPONTLSKSS-----EVFTTF 411
QY 309 EGPAWKISRKQGVITLKNNGDFFIANEGRRPIYIDGRPVLCGSKWRLSNNSVVEIASLRF 368
DB 412 KQNEKMTKK--IKLLEKETTY-----RSRWESSNKALLEMAEKET 451
QY 369 V 369
DB 452 V 452

RESULT 9
US-09-976-740-5
; Sequence 5, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289

;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5
;; LENGTH: 557
;; TYPE: PRT
;; ORGANISM: Oryctolagus cuniculus
US-09-976-740-5

Query Match 5.4%; Score 107; DB 9; Length 557;
Best Local Similarity 18.9%; Pred. No. 0.44;
Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17;

QY 16 EPSSEKKVSKA-----PSTPV-----PSPAPAGITLTKVKS-----KQPLQVT 57
DB 105 EPDAEKSRAYVANGEPGTPTPVNNGEKESKAPG--TEIRISDEVGDHRPRQEKK 163
QY 58 KDLGRWKPANDLLINAVLQT-NDLTSVHLGVKFSCTFTLREVOERWYALLYDPISTKLA 116
DB 164 KAKGLGKE-----ITLMQTLNTLSTPEEKALCKKYAELLEHRNSQOMKLLQKKQ 217
QY 117 COMRQLHPEAIAAQKALFESKA----- 150
DB 218 SOLVQE--KDLRGESKAVNLSKLSLCRELQHNHNSLKEGVQARAEERKEKREYTS 275
QY 151 TSQPTLETFTQDLHHRPDAFYLAFTAKALQAHWQ-----LMKOYLLPDQVQPLPKG 203
DB 276 HFQMTLNDIQLOMEQHNE-----RNSKLRQENMELAEKLIQYLEEHEHNI----- 323
QY 204 DQVNLFSDAEDLIDSKLDMDEVLHEHLMVADRROKREIROELHKKQVLDVSTG 263
DB 324 DKVFYGHKDLQQLVDAKLQQAQEMLKE-----AEERHQRE-----KQFLLEAVES 369
QY 264 MSSPFDNQTLAVLRGMRYLMRSREI--TLGRATKDNQIDVDLSLEGPWKISRKQGV 321
DB 370 QRMCELMKQOETHLKQOLALYTEKEFEFPONTLSKSS-----EVFTTFQOENKMTKK--I 422
QY 322 IKLKNNGDFFIANEGRRPIYIDGRPVLCGSKWRLSNNSVVEIA 364
DB 423 KLEKETTY-----RSRWESSNKALLEMA 447

RESULT 10
US-09-962-055-5
; Sequence 5, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/09/962,055
3 FILING DATE: 24-Sep-2001
4
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: 08/979,608
7 FILING DATE: 26-NOV-1997
8 APPLICATION NUMBER: US 60/031,930
9 FILING DATE: 27-NOV-1996
10
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Myers, Louis
13 REGISTRATION NUMBER: 35,965
14 REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 617/542-5070
17 TELEFAX: 617/542-8906
18
19 INFORMATION FOR SEQ ID NO: 5:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 557 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25 FRAGMENT TYPE: internal
26 SEQUENCE DESCRIPTION: SEQ ID NO: 5:
27
28 US-09-962-055-5
29
30 Query Match 5.4%; Score 107; DB 10; Length 557;
31 Best Local Similarity 18.9%; Pred. No. 0.44;
32 Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17
33
34
35 16 BPSSSEKKVKSKA-----PSTPV-----PPSPAPAGLTGRVKYS-----KQPLQVT 57
36 105 EPEDAEKSRAYARRNGEPPEPTGVNNGEKETSKAEPG-TETRTSDVGDGRHRRFOEKK 163
37 58 KDLGRWKPANDLLINAVQT-NDLTSVHGVKFSCTFLREVOERWYALLYDPVISKLA 116
38 164 KAKGIGKE-----ITLMQTLNTLSTPEBKLAALCKYVAELLEHRNSQKQKLLQKQ 217
39 117 QQAMQQLHEPAIAIOSKLFSSK-----EEOILSKVGS 150
40 218 SQTVOE--KDLRGESKSIILARSKLESLCRELQRRNSLKEGVQRAREEKEKREKVS 275
41 151 TSQPLTETQDILLHHPDAFYLAFTKALQAMQ-----LMQVYLLEDQVQGLPVG 203
42 216 HFQMTLNDIQLOMEQNE-----RNSKIQENNELAERLKLIEOYLEEHEI----- 323
43 204 DQVLFNSDAEDLIDSKLQMDREDEVLEHLMVADRQKEIRQLBQLHKQVLYVDSIG 263
44 324 DKVFFPKHDIQQLVADKALQAOEMLKE-----AEERHQE-----KDLFLKAAVSS 369
45 264 MSSPFDNQTLAVLRGWVRYLMRSHEI--TLGRATKQNDIVDLSLEGPAWKISKQGV 321
46 370 QMCELMKQOETHLKQQLALYTEKFEFPQNTLSKSS-----EVFTTFQOEMEKMTK--I 422
47 322 IYLNKNGDFEFLANEGRRPIYIDGRPVLGSSKRLSNLSVVEIA 364
48 423 KLEKETTWY-----RSRWESSKALLEMA 447
49
50 RESULT 11
51 US-10-023-529-5
52 ; Sequence 5, Application US/10023529
53 ; Patent No. US20020129388A1
54 ; GENERAL INFORMATION:
55 ; APPLICANT: Lees, Ann M.
56 ; APPLICANT: Lees, Robert S.
57 ; APPLICANT: Lees, Simon W.
58 ; APPLICANT: Arjona, Anibal A.
59 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
60 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
61 ; TITLE OF INVENTION: ATHEROSCLEROSIS
62 ; FILE REFERENCE: 10797-004001
63 ; CURRENT APPLICATION NUMBER: US/10/023,529
64 ; CURRENT FILING DATE: 2001-12-17

```

```

      PRIOR APPLICATION NUMBER: 09/616,289
      PRIOR FILING DATE: 2000-07-14
      PRIOR APPLICATION NUMBER: US 09/517,849
      PRIOR FILING DATE: 2000-03-02
      PRIOR APPLICATION NUMBER: US 08/979,608
      PRIOR FILING DATE: 1997-11-26
      PRIOR APPLICATION NUMBER: US 60/031,930
      PRIOR FILING DATE: 1996-11-27
      PRIOR APPLICATION NUMBER: US 60/048,547
      PRIOR FILING DATE: 1997-06-03
      NUMBER OF SEQ ID NOS: 53
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO: 5
      LENGTH: 557
      TYPE: PRT
      ORGANISM: Oryctolagus cuniculus
US-10-023-529-5

Query Match          5.4%; Score 107; DB 12; Length 557;
Best Local Similarity 18.9%; Pred. No. 0.44; Indels 114; Gaps 17
Matches    76; Conservative    76; Mismatches 137;

QY   16 EPPSEKKKVKSA-----PSTPV-----PPSPAPAGLTIRVKS-----KOPLOVT 57
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   105 EEEDAEKRAYANRNGEPPEPTPVNNGEKETSKARP--TEEIRTSDEVODRRHRRPQEK 163
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY   58 KDLGSRKPANDLLINAVLQT-NDITSVHLGVKFSCEFTLRVEGRWALYDPVISKLA 116
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB   164 KAKGIGKE-----ITLMQTLNTLSTBEKLAALCKKYAELLIEHRNSQKOMKLQKKQ 217
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY   117 QAMQLHPREALAQSLAFSKA-----EROLLSKVGS 150
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB   218 SOLVGE--KDRIHGSHSKATILRSKLSLCRELQHNRSLKEEGVQRAREEEKREVTIS 275
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY   151 TSQPLETEFODLIHRHPAFYLARTAKALOAWQ-----LMKOYLLLEDOTVOPLPKG 203
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB   276 HQMTLNIDIQLOMEGHNE-----RNSKLRQEMMELAEARKULIEGYELREBNI----- 323
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY   204 DQVLNPSADIEDLIDDSKLDKMDREDELEHMLVADRROKREIRQLDQELHKWOVLVDISTG 263
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB   324 DRVFGHKDLQQQLVDAKLQAQOEMLKE-----AEERHORE-----KDFLKEAVES 369
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY   264 MSSPFDNQTLAVLRGRWRYLMSREI--TLGRATKNQIDVDLSLEGPAKISRKGCV 321
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB   370 GRMCGLMQOETHLKQOALALYTEKEEFONTLSKSS----EVFTTFQOEMEKTWK--I 422
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY   322 IGLKNNGPFIANEGRRPIYIDGRPVLGSGKRWLSSNNSVEIA 364
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB   423 KLLEKETMY-----RSRWESSNKALLERA 447
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 12
US-10-023-523-5
Sequence 5, Application US/10023523
Patent No. US20020152485A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Ajiona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/10/023,523
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US/09/616,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930

```

PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 557
TYPE: PRT
ORGANISM: Oryctolagus cuniculus
US-10-023-523-5

Query Match 5.4%; Score 107; DB 12; Length 557;
Best Local Similarity 18.9%; Pred. No. 0.44;
Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17;

QY 16 BBSSEKKKVSKA-----PSTPV-----PPSPAPAGLTKRVKS-----KQPLQVT 57
DB 105 BPEDEKSRAYAVRANGPEPPTPVVNGEKETSKAPG--TEETRTSDEVDGRDRRPPQEK 163
QY 58 KDLSGKRPANDLLINAVLQTNLTSTVHLGVKFSRFTLRVQSRWYALLVDPVTSKLA 116
DB 164 KAKGKGE-----ITLMQTLNTLSTPEKLAALCKKAYELIEHRNSQKQMKLQKKQ 217
QY 117 QQAMQLPHEALIAIQSKALFSKA-----EBQLLSKYGS 150
DB 218 SQLVGE--KDLRGESHKAILARSKLESICRELQHRNSLKEEGVQRAEERKKEVTS 275
QY 151 TSQPLTFPQDLHHRPFAFYIARTAKALOAWQ-----LMKQYLLQDQTVPLRKG 203
DB 276 HFQMTLNDIQLOMEQHN-----RNSKLRQENMEELAKKLEIQEYELREHEI----- 323
QY 204 DQVLFSDAEDLIDSKLQDMDEVLEHMLVADBRQKEIRIQLOELHKQVLVDSITG 263
DB 324 DKVFYKHLQOOLVDKALQOQEMLKE-----AEERHORE-----KDFLKEAVES 369
QY 264 MSSPFDNQTLAVLRGRVYLMRSREI--TVGRATKQNDIVDSLBPAPWIKRQGV 321
DB 370 QRMCELMQOETHLKOQALALYTEKEEFQNTLSKSS-----EVFTTFQOEMKMTKK--I 422
QY 322 IXLKNGDFPIANEGRRPIYIDGRVYLCGSKRRLSNNSVVEIA 364
DB 423 KLEKETTY-----RSRWESENKALLEMA 447

RESULT 13
US-09-833-790-435
Sequence 435, Application US/09833790
Patent No. US2002006828A1
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Wang, Tongtong
APPLICANT: Secret, Heather
APPLICANT: Mohamath, Radoch
APPLICANT: Indrias, Carol Y.
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.512
CURRENT APPLICATION NUMBER: US/09/833,790
CURRENT FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 435
LENGTH: 519
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-790-435

Query Match 5.3%; Score 106; DB 10; Length 519;
Best Local Similarity 21.8%; Pred. No. 0.49;
Matches 87; Conservative 57; Mismatches 144; Indels 112; Gaps 20;
QY 62 RWKPANDLLINAVL-----QTNLTSTV--HLGVKFSRFTLRVQSRWYALLVDPVTSK 114

DB 8 RTMPAVDQLLEBALQDSPQTRSLSVFEEDAGTLTDYTNQLOAMQRYGAQNE----- 62
QY 115 LAQAMROLHDEALIAIQSKALFSKAEQQLLSKXGTSQPLTFPQDLHHRPFAFYAR 174
DB 63 -MCLATQQLSKQALAYEKQNFALGKDEEIVS-----TLH-----YFSK 100
QY 175 TAKALQ-AHQOLMKQYLLQDQTVQPLPKGDQVLFNSDAEDLIDSKLQDMDEVTL-BHE 232
DB 101 VVDELINLHTELARQ---LADTWLP-----IIQFRE-KDLTEVSTKDLFGLASNEHD 150
QY 233 LM-----VADRRQKEIRIQLO-----ELHKQVLVDSIT 262
DB 151 LSMAYSRLPKKENKXTEVGKEVAARRKQHLSSIQYCALNALQYRKQAMMEPMI 210
QY 263 GMS--SPFDNQTLAVLRGRVRL-----MRGREITLG-----RATKQNDIVDSL 308
DB 211 GFHQQINFQKGAEMFSKRMDSFLSVADWQSIQVELAEAEKMRVSOQELLSDVESV 270
QY 309 EGP-----AMKISR-----KQGVIKLKNNGDFPIANEGRRPIYIDGRVYLC-----GS 351
DB 271 YTPSDVAAPQINRNLQKAGYINLRNKTGLVTTTWERLYFTTQGNLMQPRGAVAGL 330
QY 352 KWRLSNNSV--VEIASLRFVELI--NODLIALIAEBAK 386
DB 331 IQDLNCGSVMAVDCEDRRYCFQITTPNGKSGIILQAESRK 370

RESULT 14
US-09-864-761-38419
Sequence 38419, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomic-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38419
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049872.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.72
; OTHER INFORMATION: EST HUMAN HIT: AU138211.1, EVALUATE 2.00e-83
; OTHER INFORMATION: SWISSPROT HIT: Q10411, EVALUATE 3.00e-18
US-09-864-761-38419

```

```

Query Match 5.1%; Score 102.5; DB 10; Length 751;
Best Local Similarity 20.4%; Pred. No. 1.7;
Matches 84; Conservative 70; Mismatches 126; Indels 129; Gaps 19;

```

```

QY 18 SSSSEKKKYSKAPSTPPSPAPAPGCTTKR-----VKKSKQPLQYTKD-LGRMKP 65
DB 237 NSEKNDLSKASTERL-----VKGIKERLEIKLNEKNISLTKQIDQLSKDEVGK--- 287
QY 66 ANDLLINAVLQNDLTSVHLGVKFSCEFTLREVV-----QBRWYALVLPVLSKLACQAM 120
DB 288 -----LTQITQOKDEIGALHARISSTHTQVUVYLOOLQYANERKRAVLNEMKTR 341
QY 121 ROLHREA-----IAIISKALFSKAE--OLSKVSGTSQPTL-ETFODL-LHRHPD 168
DB 342 ESHKTEVHKMDIVAAKEALIKLODENKKLSTFESSGQDMFRETQJNSRIIREKD 401
QY 169 AFY--LAPAKALQAHWOLMKOYLLLEDQTVORLPKGDV--LNSDAEDLIDD-SKLKD 223
DB 402 IEIDLQKQOTLLA-----VLQTSSTGNAGGVNSQNPBELLQERDKLQ 447
QY 224 M--RDEVLEHEIMVADRQREIRLOEELHKW--QVLVDSITGMSSPDFDQTLAVLRG 279
DB 448 QVKKEWKOQVTVYQNMQHSQAQLOEBLHQLOAVLVDS-----DNNSKL----- 494
QY 280 RMVRYLMSREITLGRATKQNDIVDLSLEGPAWKISRQGVYKLKNGDPFLANEGRRP 339
DB 495 -----QVDYTGILQ-----SYQNETKCLKN----- 514
QY 340 IYIDGRPVLCGSKWRLSNNSVVEIASLRFVFLINDLILIRAAKTIPTQ 390
DB 515 -----FGQELAQVGHISIGLCNTKDLILKGLDIIPQSSASLTPQ 556

```

```

RESULT 15
US-09-815-242-11259
; Sequence 11259, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zysek, Daniel W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard

```

```

; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11259
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11259

```

```

Query Match 5.0%; Score 100.5; DB 10; Length 475;
Best Local Similarity 19.1%; Pred. No. 1.3;
Matches 84; Conservative 62; Mismatches 136; Indels 157; Gaps 18;

```

```

QY 14 GSEBSSSEKKKYSKAPSTPPSPAPAPGCTTKR-----VKKSKQPLQYTKD----- 61
DB 7 GLSTANTENKXAKAPKT-----AYVNDGAEFSRWGQCSA 44
QY 62 --RMKRPANDLLINAVLQNDLTSVHLGVKFSCEFTLREVV-----QBRWYALVLPVLSKLACQAM 119
DB 45 CKAANTTTEVRLISAKSKNDPFGSAGETQAKIOTLSEI----- 84
QY 120 MCOLHREAIAIISKALFSKAEOLSKVSGTSQPTLFTFODL-LHRHPDAFYLAFTAKAL 179
DB 85 -----SLQETPRFSSGFSSELDRLVGGGIVGSAI--LIGHPCA--GKSTILL 128
QY 180 QAHNOLMKQ--YLLLEQTVORLPKGDV--LNSDAEDLIDD-SKLKD 223
DB 129 QVMGSLAKMTALVYTGSESLQOVAMRASRLG-----LPBDQKMLSETSVBOICNLA 181
QY 237 DRQKREIRLOEOLHKWQVLVDSITGMSSPDFDQTLAVLRG-----MVRYLMSREI 291
DB 182 D-----QKRPQI-----IYVDSIQVHMLADIGSSPSSVAGVRECAFLTRY- AKTRQV 228
QY 292 TL--GRATKD-----NQIDVDLSLEGPA--WKISR----- 317
DB 229 AIIWGHVTKDGTLAGPKVLEHAIDCSLLLEGADSRYTLRSHKRRGAVNELGVFGWT 288
QY 318 KQGVTKLKNQDPFLANEGRRPIYIDGRPVLCGSK-----KLSNNSVVEIASL----- 366
DB 289 EQLAEVKNPSAIFLSR-----GDEITSGSSVMVWLGSTRPLVLEIQALVDHSMILA 339
QY 367 ---RFVFLINDLILAIRA 382
DB 340 NRRVAVGLEQNRALDLIA 358

```

```

Search completed: March 27, 2003, 05:29:56
Job time : 34 secs

```

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:15:30 ; Search time 41 Seconds
(without alignments)
914.449 Million cell updates/sec

Title: US-09-744-125A-2

Perfect score: 1996
Sequence: 1 NSARGVEPGRCGSGSEPPSSS.....LINODLIATRAAKITPQ 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	231.5	11.6	408	2 T23127	hypothetical prote
2	205	10.3	735	2 T47594	hypothetical prote
3	142	7.1	551	2 D96632	hypothetical prote
4	116.5	5.8	642	2 T39607	fork head protein
5	116	5.8	676	2 AE2417	hypothetical prote
6	116	5.8	2346	2 T33829	Tpr homolog - frui
7	115	5.8	699	1 S38982	kinesin-related pr
8	111	5.6	804	2 A96494	protein F7F2.16 [i
9	106.5	5.3	1140	2 G89113	plectin - human
10	106.5	5.3	4574	2 G02520	plectin [imported]
11	106.5	5.3	3684	2 A59404	dyctrophin, muscle
12	106	5.3	3685	1 A27605	dyctrophin, heavy
13	104.5	5.2	4588	2 T28667	dyctrophin, heavy
14	103.5	5.2	558	2 G96522	actin binding prot
15	103.5	5.2	1825	2 T42725	actin binding prot
16	103.5	5.2	1885	2 T30847	actin binding prot
17	103.5	5.2	2033	2 T30849	actin binding prot
18	103	5.2	810	2 B70154	DNA gyrase chain A
19	102	5.1	411	2 T25223	hypothetical prote
20	102	5.1	968	2 T51523	clpB heat shock pr
21	102	5.1	3678	2 S28916	dyctrophin - mouse
22	101.5	5.1	1109	1 B45345	HIV-1 retropepsin
23	101	5.1	943	2 S68824	rngB protein, cyto
24	100.5	5.0	475	1 T64131	DNA repair protein
25	100.5	5.0	853	2 T10591	hypothetical prote
26	100	5.0	407	2 T08496	probable transcript
27	100	5.0	572	2 H86257	protein F501.2 [i
28	100	5.0	859	2 T01461	hypothetical prote
29	100	5.0	971	2 T10678	hypothetical prote

30	100	5.0	2223	2 S65074	pyrimidine synthes
31	100	5.0	15281	2 S41309	cyclosporin synthe
32	99.5	5.0	417	2 E64200	serine-tRNA ligase
33	99.5	5.0	496	2 E47199	1-aminocyclopropan
34	99.5	5.0	911	2 F96771	heat shock protein
35	99.5	5.0	1490	2 S32373	DNA-binding protei
36	99.5	5.0	2068	2 A47371	transcription init
37	99	5.0	2244	2 T11616	carbamoyl-phosphat
38	98.5	4.9	1205	2 D83862	hypothetical prote
39	98	4.9	743	2 H96483	protein F7F2.18 [
40	98	4.9	1278	2 T27925	hypothetical prote
41	97.5	4.9	677	2 H64574	DNA topoisomerase
42	97.5	4.9	835	2 AD2441	endopeptidase Clp
43	97.5	4.9	958	2 T10679	hypothetical prote
44	97.5	4.9	1048	1 BVESC3	exonuclease (EC 3.
45	97.5	4.9	4644	1 A38905	dynein heavy chain

ALIGNMENTS

RESULT 1

T23127
hypothetical protein H28016.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23127

R:White, S.
submitted to the EMBL Data Library, June 1998

A:Reference number: Z19688

A:Accession: T23127

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-408 <MIL>

A:Cross-references: EMBL:AL023815, PIRN:CAA19428.1, GSPDB:GN00019, CBSP:H28016.2

A:Experimental source: clone H28016

C:Genetics:

A:Gene: CBSP:H28016.2

A:Map position: 1

A:Introns: 39/2; 62/3; 89/3; 254/3; 268/2; 321/2; 402/2

Query Match 11.6%; Score 231.5; DB 2; Length 408;
Best Local Similarity 21.6%; Pred. No. 7.7e-09;
Matches 87; Conservative 69; Mismatches 134; Indels 113; Gaps 12;

QY	9	PERCGSPSSSEKKVSKAPSTP-----VPEPPAPPGUTK--YKSKQPLQVTKD 59
DB	44	PNEISGFKKQKIEENSEIHAETESVEAOWMKKPEEPBLKRRKILIHNEPSITETNP 103
QY	60	LGRMKPANDLLINAVLOTNDLTSVHLGVKFCRTLAEGVERWYALLYDPVISLACQA 119
DB	104	KRIWTVADLLALTVAVHOCRTFHSNLTFSRKTFSDVEERYCQMLDEIISKAKVR 163
QY	120	MRQLPPEAIIAIOAKALFSKAE----- 142
DB	164	LDAMHRLKAQIEATPTTRNEERTIMELAENOLKNARKERQVENHTVLTILHKKILD 223
QY	143	-----QLSKVGSSTGPT-----LETFODLLHRHPDAFYLAATKALQA 181
DB	224	GNRYGSKFSFAKICLAELIVQIGSTQAKKSGLFQKQYQ-LFEKERTQSFHKSRTPQVLS 282
QY	182	HMQLMKQVYLLLEDQVQPLPKGDQVLAIFS-DAEDLIDSKLQKMDVLEHLMVADBRQ 240
DB	283	HYRRITGY-----RSEGNNNNYMWQLLEALTGNTWDFD--INAPLQARSRY 327
QY	241	KREIRQ--LEQELHKQVYLVDSITGMSPPDFNQTLAVLRGMYVYLRSEIITLGRATK 298
DB	328	AAISRRPLSLGILNRFK-----TSGSVDP--NAIMINIQFLQYANTGKSVTMGRASL 378
QY	299	DNQIDVDVLSLEGPAWKISRKQGVIKLKNQDFFIANEGRRPIY 341
DB	379	NEKIDIDLSKSGPATKVL-----EGRROLF 403

RESULT 2

hypothetical protein T1ZE18.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T17594
R:Bloembergen, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24469
A:Accession: T47594
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-735 <BLO>
A:Cross-references: EMBL:AL132971
A:Experimental source: cultivar Columbia; BAC clone T1ZE18
C:Genetics:
A:Map position: 3
A:introns: 27/2, 56/3, 570/3, 588/1, 637/3, 666/3, 709/3
A:Note: T1ZE18.40

Query Match	10.3%;	Score 205;	DB 2;	Length 735;
Best Local Similarity	32.6%;	Pred. No. 1.3e-06;		
Matches	56;	Conservative	35;	Mismatches 63;
			Indels	18;
			Gaps	4

QY 208 NFSDAEDLIDDSK-L-DNRD---EVLHELMVADROROREIROLEQELHKQVLVDSIT 26

```
Db      563 NYSDEAMILMDLEPPDDQDNFDFLEVSKYQ---SQDMKRTIRLEQAASHSYMQRATASR 61
```

QY 263 GMSPPDNQTLAVLRGMVRYLMRSREITLGRATKDNQIDVDLSLEGPAWKISRKQGI 322

```

Db 619 G-----AFAVLGYRYSKHYIKKEPEVLVGRSTEDLAVDIDLGREKRGSKISRQAI 66
      |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

```

QY 323 KLKNGDFIANEGRPYIDGRPVLCGSKWRLSNNSVVEIASLRVELINQ 374

```

Db      670  RLGGDSFHIKNLGKYSISVNEKEVDPGSLILKSDCLVEIRGMPIFETNQ 7211
          ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

RESULT 3

hypothetical protein F8A5.22 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96632
R:Thelloglis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Native 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Saldegg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96632
A:Status: preliminary
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <STO>
A:Cross-references: GB:AE005173; NID:g2462740; PIDN:AA871959.1; GSPDB:GN00141
C:Genetics:
A:Gene: F8A5.22
A:Map position: 1

Query Match	7.1%;	Score 142;	DB 2;	Length 551;
Best Local Similarity	23.5%;	Pred. No. 0.025;		
Matches 47;	Conservative 38;	Mismatches 73;	Indels 42;	Gaps 4

211 DAEDLID-DSK-L-----DMRDEVLEHELMVADRQKREIRQLEQELHKQVLDVSTG 26

```
Db      339 DGEEDIDAMIRKLNLPDDSDSCFNREEMMSKHPRHALGLEQ-----CTRTS 380
```

QY 264 MSSPFDNQTLAVLRGMVRYLMSREITLGRATKDNQIDVDLSLEGPAWKISRKQGVK 322

```

Db      |      : :|||      : :| ||| : ||| :      : ||| : ||| :| :|
390 MGRAIMFHGAI AVLHCPDSKH FVRKREVI IGRSSGGLNVDIDL GKYNYSKISRQALVK 445

```

QY 324 LKNGDFFIANEGRRPIYIDGRPVLCGSKRRLSNNS----- 359

Db 450 L E N Y G S F S L K N L G K O H I L V N G G K L D R G Q I V T L T S C S S I N V S T S L C F V Q V C F C I F L S T R T 509

QY 360 --VVEIASLRFVFLINQDLI 377

```

Db      510 A Q I L Q R G I T F V E K I N K E A V 529
          ::| : ||| : : :

```

RESULT 4

fork head protein type transcription factor - fission yeast (*Schizosaccharomyces pombe*)
 C1:Species: Schizosaccharomyces pombe
 C1:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C1:Accession: J39607
 R1:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, May 1998

```
;Accession: T39607
;Status: preliminary; translated from GB/EMBL/DBJ
;Molecule type: DNA
```

A; Cross-references: EMBL:AL023554; PIDN:CAA19034.1; GSPDB:GN00067; SPDB:SPBC16G5.15C
Experimental source: strain 872b-1; Cosmid C16G5

A/Experimental Source: Brian Smith, Cornell Univ
C/Genetics:

```

A:Map position: 2
A:Tab: 171/0

```

A/INCLIONS: 1/1/2

Best Local Similarity 24.0%; Pred. No. 1.9;

Matches	43;	Conservative	42;	Mismatches	11;	Indices	11;	Gaps	
---------	-----	--------------	-----	------------	-----	---------	-----	------	--

RESULT 5

hypothetical protein alr4893 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Name: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AE2417
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807, MUID:21595285, PMID:11759460
 A:Accession: AE2417
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-676 <KIR>
 A:Cross-references: GB:BA000019; PIDN:BA076592.1; PID:gi1134030; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 C:Gene: alr4893 -

Query Match 5.8%; Score 116; DB 2; Length 676;
Best Local Similarity 19.8%; Pred. No. 2.2;
Matches 69; Conservative 64; Mismatches 126; Indels 90; Gaps 14.

```

Qy 5 GVEPGRCSGSESPSSSEKKVSK-----APSTP---VPPSPAPAGLTGRVKSQKPLQV 56
Db 53 GNVTPPEPTEVTPPQSTKSRVIAQNSPVLPSPTRPVPPAPRATVNE-----LVV 104
Qy 57 TKDLGRMKRPANDLLINAVLQTNLTSVHLGVKFSRFTLREVGQRMWALLVDPVSKL- 115
Db 105 T-----ATDVQIVAGTPPELQEIIRQVVKIKQGTGDSIQFQDLQDVALLLETGTFASAN 156
Qy 116 -----ACQAMROLHEAIAAIG---SKAL-FSKAEQLSKVGSQPTLETFQDL 163
Db 157 VNSRTTSGLVNVVQVQVPIVASLQLTGAKALTVSAQPRFQSGIKGRISP-----EGL- 210
Qy 164 HRHPDAFLATRAKALQAMQOLMQLKQYLLDQTVPLPKGQDVLFNSDAEDLIDSKLKD 223
Db 211 -----KQAVQAVNQWYADNGYNLARVLSIFPNQG--ILNINVAEGVLSIKRFR 258
Qy 224 MRDEVLEHLMVADRQKREIRQLEQELHKQVLDVS-----ITGMSPPDFNQTLAVLRG 279
Db 259 VNDD-----GKTIDSNQNPVGGRTPDKPDLRQQLKLPQ 291
Qy 280 RMVRYLMSREITTLGRATKQNDIDVLSLEGPAWKISKQGVTKKANG 328
Db 292 QVFQENIVKQDVQQLYRTGLPQ-SYVNAFAGDATTLDL---IYELKENG 336

RESULT 6
Tpr homolog - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C/Accession: T13829
R/Zimovska, G.; Aris, J.P.; Paddy, M.R.
J. Cell Sci. 110, 927-944, 1997
A/Title: A Drosophila Tpr protein homolog is localized both in the extrachromosomal char
A/Reference number: 217786; MUID:97296455; PMID:9152019
A/Accession: T13829
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-2346 <ZIM>
C/Genetics:
A/Cross-references: EMBL:U91980; NID:91923273; PID:91923274; PIDN:AA047506.1
A/Map position: 2R

Query Match 5.8%; Score 116; DB 2; Length 2346;
Best Local Similarity 23.3%; Pred. No. 13;
Matches 86; Conservative 55; Mismatches 126; Indels 102; Gaps 21;

Qy 45 KRVKSKQPLQVTKDLGRMKRPANDLLINAVLQTNLTSVHLGVKFSRFTLREVGQRMW 104
Db 340 KKIQMEQELASANDL-LKQARESNLSAICQLAPSAV-----ASLIRSDSLTEL 391
Qy 105 ALLYDPVSKLACQAMROLHEAIAAIGSKALFSAEBO--LISKVGSQPTLETFQDL 162
Db 392 YSMYAKSSEEL--EMRNCETIQL-KLQKSIIAEISEAPFLKQNSPYQMKETNSL 447
Qy 163 LHRPDAFY-----LARTAKAL-----QAHMQLMKQYLLDQ----- 195
Db 448 LREHDELLQNKCLERELERALSTLNHQNENKQLQHTDLSRQVCMILDELINCRAGV 507
Qy 196 ---TYQP---LPK-----GDQVLFNSDAEDLID-DSKLKMDVLEHLMVADR- 239
Db 508 KHVRIQTPRQLTSSLSIDNLVTFSSIEELVDNRVYLLNMSRELT--ELLEASERKNDK 565
Qy 240 -----QKEIRIQLEQELHKQVLDVSITGMSPPDFNQTLAVLRGRWRY--LMSREIT 292
Db 566 MLLEBSKNIHRLDA--RFALEDLITQ-----KNNVTITLISKCRKYKLYAAQCK 616
Qy 293 LGRATKD-----NQIDVDLSLEQPA--WKISRK-----QGVIK---LKN 326
Db 617 LGQNTVLDLDDSNLEPNDSLDTS--EQPANPEKSRKLEKRVQLRQLEGEVKKYASLKE 675

```

```

Qy 327 NGDFPIANE 335
Db 676 NYDYTSEK 684

RESULT 7
Kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)
N/Alternate names: Kinesin-2 chain A; KRP (85/95) 85K chain
C/Species: Strongylocentrotus purpuratus (purple urchin)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
R/Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.
Nature 366, 268-270, 1993
A/Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.
A/Reference number: S38982; MUID:94050179; PMID:8232586
A/Accession: S38982
A/Molecule type: mRNA
A/Residues: 1-699 <COL1>
A/Cross-references: EMBL:L16993; NID:9295245; PIDN:AAA16098.1; PID:9295246
A/Accession: S72551
A/Molecule type: protein
A/Residues: 2-5, 'X', '7-11, 59-64, 125-132, 222-226, 'X', 228-230 <COL2>
C/Complex: heterotrimer of a 115K chain and two kinesin-related chains of 95K (PIR:S5869
C/Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
C/Keywords: Arp; heterotrimer; microtubule binding; nucleotide binding; P-loop
F:97-104/Region: nucleotide-binding motif A (P-loop)
F:103/Binding site: ATP (Lys) #status predicted

Query Match 5.8%; Score 115; DB 1; Length 699;
Best Local Similarity 20.6%; Pred. No. 2.8;
Matches 82; Conservative 58; Mismatches 138; Indels 120; Gaps 16;

Qy 6 GVEPGRCSGSESPSSSE-----KKVSKASPFPVPPSPAPAGLTGRVKSQKPLQV 57
Db 377 GLDDDESGSESGDEBAGGVKKRKGKPKRKL--SPETIMAMQKIDEEKALEK 434
Qy 58 KDLGRMKRPANDLLINAVLQTNLTSVHLGVKFSRFTLREVGQRMWALLVDPVSKLAC 117
Db 435 KD-----MVEEDRNTVA-----RELQRR-----ESELHK--A 459
Qy 118 QAMROLHEAIAAIGSKA-----LFSKAEQLSKVGSQPTLETFQDLHRPDAFY 171
Db 460 QDDQKILNEKLAIAIOKLIIVGVVDLLAKSBO-----EQLEQ----- 497
Qy 172 LARTAKALQAMQOLMQLKQYLLDQTVPLPKGQDVLFNSDAEDLID-DSKLKMDRDEV 229
Db 498 -----SALEMKERMAKQSMKRMKEERQEMDIEKYSLSQDEAHGKTKLKKVMTLM 552
Qy 220 EHELMVADRQKRE-----IRQLEQELHKQVLDVSITGMSPPD-----NOTL 274
Db 553 QAKSEVADMOAHEQREEMALLENVARELSREURLSLIIDSFIPOEFQEMIDQYVMMNDI 612
Qy 275 AVLGRWRYLMSREITTLGRATKQNDI---DVDLS-----LEGPAWKISKQGVTKL 324
Db 613 GBWQKCAVYIGNNRKKQTPVADKDKSLAYGEADSNVFLTYNLBGGMKTKPSQG----- 668
Qy 325 KNNGDFPIANEGRPYIDGRPVVLGSKRWRLSNNSVE 362
Db 669 -----KSGRPXTSSGRP-KTGKKKQASMASSID 695

RESULT 8
protein F7P22.16 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: A96494
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000

```

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzman, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:21016719; PMID:11130712
A:Accession: A96494
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-804 <STO>
A:Cross-references: GB:AEO05173; NID:g6691192; PIDN:AAF24530.1; GSPDB:GN00141
C:Genetics:
A:Gene: F7F22.16
A:Map position: 1

Query Match 5.6%; Score 111; DB 2; Length 804;
Best Local Similarity 23.0%; Pred. No. 6.4; Indels 60; Gaps 11;
Matches 67; Conservative 41; Mismatches 123;

Qy 97 REVERRYAL---LYDPVISKLACQAMROLH-PEAIIAQSKALFSKAEOLLSKVGS 152
Db :|::||::||::||::||::||::||::||::||::||::||::||::||
47 RMATERKEILDEVEDEMEPQTRATKLHKPDLLPTEBYRLFLNFECSTRYCASTS 106
Qy 153 QPTLETFODLLHRHPDAFYARAKALOHWOLMKOXYLLEDQTVLPKGDVLNPSDA 212
Db |::||::||::||::||::||::||::||::||::||::||::||::||
107 LAQLGLLEDEVQH-----LYQSCHELDITLMAY-----PYVAVEBETIOPLTLQVKLYQGWT 156
Qy 213 EDLIDDSKLDMDREYLVEHELMVADR-----QKRERLOELHKKWQ 255
Db ||::||::||::||::||::||::||::||::||::||::||::||::||
157 SDELDCGELGFRLRSVGHERRSLTKRLGLEFGPSGTSGSKPKYEEDLKDL-----WI 210
Qy 256 VLVDISIRGMSSPPDNQNTLAIVLRGRMVRVLMR-----SRKITGRATKDNDIVDLSL 308
Db 211 TLGGSVPLANSRSNSQ----IRSPVIRYFQRSVANVLXSRETT-GTYVNSDMEMTAMAL 265
Qy 309 EGPAWKISRKGQVIKLNNGDFFIANEGRRPIYIDGRPVLCGSK-WRLSN 358
Db 266 -----KGTLRQTNGMSLGCKVNNTPLSILLIHLGCVKSWAVSN 306

RESULT 9
G89113
protein ZK742.1 [imported] - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C.Accession: G89113
R.anonymus, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A.Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A.Reference number: A75000; PMID:99069613; PMID:8851916
A.Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A.Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G89113
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1140 <STO>
A:Cross-references: GB:chr_V; PIDN:AB04981.1; PID:g1465818; GSPDB:GN00023; CESP:ZK742.1.L.1
A:Map position: 5
C:Superfamily: Arabidopsis thaliana exportin 1

Query Match 5.3%; Score 106.5; DB 2; Length 1140;
Best Local Similarity 21.4%; Pred. No. 22; Indels 75; Gaps 18;
Matches 79; Conservative 61; Mismatches 155;

Qy 50 SKQPLQVT-KDIGR---WKPANDLLIN--AVIQTNLT-SVHLGVKFSGCRFTLREVQR 102
Db :|::||::||::||::||::||::||::||::||::||::||::||::||
217 SNENQAATKTIGLFPLTWIPGYVFETINTELLSENFLELVYRVVALQCTEISQIVE 276
Qy 103 WYALLYPVSIKLACQAMROLHPEAIIAQSKALFSKAEOLLSKVGSQPTLETFODL 162
Db :|::||::||::||::||::||::||::||::||::||::||::||::||

```

Db      277  TNDPSIDELKVKFCSTMRHISQVLSDLDLAAYKKAASDDQKLISSLAQVLVAFIKH 336
Qy      163  LHRHPDAFYLTAKAL--QAH---WQMKQYVLLDQTVPLPKGDVLFNSDAED--- 214
Db      337  VHLIEVTDEPLTEAKILMRSHDYAIQLLKLTLIEEMEV-----FKVCLDWC 385
Qy      215  -----LDDSKLKMREDEVLEHMLVADRQ--KEIRQLQELHKWQVLYVDSI 261
Db      386  WLTAELYRICPFLQPSITLVGMSQVREH---PRRQLYRYLSQLRSTM-----I 431
Qy      262  TGMSSPDPNQTALV--RGRMYLYMRSRE-ITLGRATKNOIDVDLSLEGPAWKISRK 318
Db      432  SRRAKE---EVLIVENQGEVREWKDIDSIALYMRRTLYVL-THLDKQDEYVKWT 487
Qy      319  QGVYIKLKNNGDFPIANERRPPIYIDGRPVLCGSKWRIS--NNSVVEIASLRVFLINQDL 376
Db      488  EKLSQVNGGEFSWKMLNR-----LC---WAVGSIGTWEDEBEKRFVLVIRDL 534
Qy      377  IALIRAEAAK 386
Db      535  LGLCEQKRGK 544

RESULT 10
G02520
pIelectin - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Aug-1999
C:Accession: G02520
R:McLean, W.H.I., Smith, F.J.D.
submitted to the EMBL Data Library, March 1996
A:Reference number: H01385
A:Accession: G02520
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4574 <MC>
A:Cross-References: EMBL:U53204; NID:g1477645; PIDN:AA05427.1; PID:g1477646
C:Gene: PLECI
C:Genetics:
C:Superfamily: pIelectin; alpha-actinin actin-binding domain homology; ribosomal protein S
F:68-283/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match 5.3%; Score 106.5; DB 2; Length 4574;
Best Local Similarity 23.5%; Pred. No. 1.5e+02;
Matches 70; Conservative 52; Mismatches 119; Indels 57; Gaps 13;

Qy      122  QLRPEA-IAIISKALPESKAEOLISKVGSTOPTLETFODLHRHPDAFYLTAKALQ 180
Db      2488  KIQQEKALQLKSEWQTVQEQQLQETQALQGSFLSKDSILQR--ERFIQEKAKLEQ 2545
Qy      181  AHW-QLMKQYVLLDQTVQPLPKGDVLFNSDADLDIDSXLKDMRDEVLEHMLVADR 239
Db      2546  LPQDEYAKAQQLLEEQRO-----QQQMEGERQLV--ASMEEARRR--QHEAEGVRR 2595
Qy      240  QKRREIQLQELHKWQVLYVDSITGMSPPDPNQ---TLAVLRGMRYLMSRREITLGR 295
Db      2596  KQRELQQLQEQRRQOEELAE-----ENQRLEQLQLEEGHRAALAHSEVYASQ 2646
Qy      296  --ATKNOIDVDLSLEGPAK-----ISRKQGVILKKNNGDFPIANEGRRPI---- 340
Db      2647  VAAIKTLPLNGRD-ALDPSAEAEPEHSPFDGIRRYKSAORLQAG--ILSHEELQRLAQGHT 2704
Qy      341  -----YIDGRPVLCGSKWRISNNSVVEIASLRVFLINQDLIALIRAEAA 385
Db      2705  TYDELAREDEVRYLQGRSSISAGLLKATNEKLSVYALQQLSPGALLILEAQAA 2762

RESULT 11
A59404
pIelectin [imported] - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
C:Accession: C59404; A59404

```


R;Ilu, C.G.; Maercker, C.; Castanon, M.J.; Hauptmann, R.; Wiche, G.
 Proc. Natl. Acad. Sci. U.S.A. 93, 4278-4283, 1996
 A:Title: Human plectin: organization of the gene, sequence analysis, and chromosome loca
 A:Reference number: C59404; MUID:96210632; PMID:8633055
 A:Accession: C59404
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-4684 <STO>
 A:Cross-references: GB:CAA91196; NID:g1296662; PIDN:CAA91196.1
 C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S

Query Match 5.3%; Score 106.5; DB 2; Length 4684;
 Best Local Similarity 23.5%; Pred. No. 1.6e+02;
 Matches 70; Conservative 52; Mismatches 119; Indels 57; Gaps 13;

Qy 122 QLHPRA-1AAIUSKALFKAEHQLSKYGSTQPTLFETFDLIHHPDAFYIARAKALQ 180
 Db 2598 KLQGEAKLQKSEMQVVOOEQLQETQALQOSFLSEKDSLQF--BRFIEQEKAKLEQ 2655
 Qy 181 AHM-QLMQVYLLEPQVPLPKGQVLFNFSDAEDLIDDSKDKMRDEVLHEHLMVADR 239
 Db 2656 LFOEVAKAQQLREEQNQ-----QQMEGRKQLV--ASNEARRR--QHEAEQVRR 2705
 Qy 240 QKREIRQLEOEELHKQVIVDSITGMSSPDFNQ---TLAVLRGMVRYLMSREITLGR 295
 Db 2706 KQELQQLQEQRRQOEELAE-----ENQRLREQLQLEQGHRAALAHSEEVTAHQ 2756
 Qy 296 --ATKQNDIVDLSEGGAMK-----ISRKQGVIKLKNQPFPIANERRP1---- 340
 Db 2757 VAATKTLPLNGRD-ALDGPAAAEPEHSPFDGRKRYSAQRLQAG--ILSAEELQRLAQGHT 2814
 Qy 341 -----YIDGRPVLCGSKMRLSNNSVVEIASRFVFLINODLILIRAEAA 385
 Db 2815 TVDELAREDEVRAHYLGSSISIGLLIKATNEKLSVYALQRLSPGTRALITLEKQAA 2872

RESULT 12
 A27605
 N:Alternate names: Duchenne muscular dystrophy protein
 C:Species: Homo sapiens (man)
 C>Date: 19-Nov-1988 #sequence revision 27-Jun-1994 #text change 16-Jun-2000
 C:Accession: A27605; S07710; A27162; S05291; A40134; S06051; S10346; S02242; S02
 R:Koenig, M.; Monaco, A.P.; Kunkel, L.M.
 Cell 53, 219-228, 1988
 A:Title: The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein.
 A:Reference number: A27605; MUID:88194521; PMID:3282674
 A:Accession: A27605
 A:Molecule type: mRNA
 A:Residues: 1-3685 <KOE>
 A:Cross-references: GB:M18533; NID:g181856; PIDN:AA53189.1; PID:g181857
 R:Rosenthal, A.; Speer, A.; Billwiltz, H.; Cross, G.S.; Forrest, S.M.; Davies, K.E.
 Nucleic Acids Res. 17, 5391, 1989
 A:Title: Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus
 A:Reference number: S07710; MUID:89345106; PMID:266885
 A:Accession: S07710
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-132; 'P', 134-622, 'I', 624-783, 'G', 785-1196, 'F', 1198-1376, 'N', 1378-1468, 'Q', 1
 A:Cross-references: EMBL:X14298; NID:g30845; PIDN:CAA34479.1; PID:g30846
 A:Note: this sequence was submitted to the EMBL Data Library, February 1989
 R:Koenig, M.; Hoffman, E.P.; Bettelson, C.J.; Monaco, A.P.; Feener, C.; Kunkel, L.M.
 Cell 50, 509-517, 1987
 A:Title: Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary
 A:Reference number: A90897; MUID:87273512; PMID:3607877
 A:Accession: A27162
 A:Molecule type: mRNA
 A:Residues: 1-497 <KO2>
 A:Cross-references: GB:M18533
 R:Cross, G.S.; Speer, A.; Rosenthal, A.; Forrest, S.M.; Smith, T.J.; Edwards, Y.; Flint,
 EMBL J. 6, 3277-3283, 1987
 A:Title: Deletions of fetal and adult muscle cDNA in Duchenne and Becker muscular dystro
 A:Reference number: S01263; MUID:8811512; PMID:3428261

A:Accession: S05291
 A:Molecule type: mRNA
 A:Residues: 404-556; 'T', 558-610, 'K', 612-622, 'I', 624-664, 'M', 665-783, 'G', 785-1137, 'PN' <C
 A:Cross-references: EMBL:X06178
 A:Note: 475-Ile and 529-Glu were also found
 R:Hoffman, E.P.; Monaco, A.P.; Feener, C.C.; Kunkel, L.M.
 Science 238, 347-350, 1987
 A:Title: Conservation of the Duchenne muscular dystrophy gene in mice and humans.
 A:Reference number: A40134; MUID:88018015; PMID:3659917
 A:Accession: A40134
 A:Molecule type: mRNA
 A:Residues: 491-1207 <HOF>
 A:Cross-references: GB:M18533
 R:Blonden, L.A.J.; den Dunnen, J.T.; van Paassen, H.M.B.; Wapenaar, M.C.; Grootscholten,
 Nucleic Acids Res. 17, 5611-5621, 1989
 A:Title: High resolution deletion breakpoint mapping in the DMD gene by whole cosmid hyb
 A:Reference number: S06051; MUID:8934515; PMID:2569720
 A:Accession: S06051
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 2147-2204 <BLO>
 A:Cross-references: EMBL:X15495; NID:g30829; PIDN:CAA33518.1; PID:g1335049
 R:Speer, A.; Billwiltz, H.; Hutch, A.; Couelle, C.; England, S.; Love, D.; Davies, K.E.
 submitted to the EMBL Data Library, February 1990
 A:Reference number: S10346
 A:Accession: S10346
 A:Molecule type: DNA
 A:Residues: 2438-2480 <SPE>
 A:Cross-references: EMBL:X51934
 R:Chamberlain, J.S.; Gibbs, R.A.; Ranier, J.E.; Nguyen, P.N.; Caskey, C.T.
 Nucleic Acids Res. 16, 11141-11156, 1988
 A:Title: Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA a
 A:Reference number: S02109; MUID:8908552; PMID:3205741
 A:Accession: S02109
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 665-722 <CHA>
 A:Cross-references: EMBL:X13045; NID:g30825; PIDN:CAA31451.1; PID:g1335048
 A:Accession: S02242
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 2098-2146 <CH2>
 A:Cross-references: EMBL:X13047; NID:g30827; PIDN:CAA31452.1; PID:g809549
 A:Accession: S02244
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 2147-2204 <CH3>
 A:Cross-references: EMBL:X13048; NID:g30833; PIDN:CAA31454.1; PID:g1335051
 A:Accession: S02109
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 2305-2365, 'K' <CH4>
 A:Cross-references: EMBL:X13047; NID:g30831; PIDN:CAA31453.1; PID:g1335050
 R:Ginjaar, I.H.B.; van Paassen, H.M.B.; den Dunnen, J.T.; van Ommen, G.G.J.B.
 submitted to the EMBL Data Library, March 1992
 A:Description: Sequence of Duchenne muscular dystrophy gene exon 60, located directly 5'
 A:Reference number: S23736
 A:Accession: S23736
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 2980-2995, 'K', 2997-3028 <GIN>
 A:Cross-references: EMBL:Z11860
 R:Heilig, R.; Lemaire, C.; Mandel, J.L.
 Nucleic Acids Res. 15, 9129-9142, 1987
 A:Title: A 230bp cosmid walk in the Duchenne muscular dystrophy gene: detection of a con
 A:Reference number: S09071; MUID:88067745; PMID:2825128
 A:Accession: S09071
 A:Molecule type: DNA
 A:Residues: SGGHSWTHCSLYRPLTL', 218-277 <HEI>
 A:Cross-references: EMBL:X06293; EMBL:Y00494
 A:Note: sequence N-terminal of residue 218 correspond to a putative exon
 R:Roberts, R.G.; Coffey, A.J.; Bobrow, M.; Bentley, D.R.
 Genomics 16, 536-538, 1993

A:Title: Exon structure of the human dystrophin gene.
 A:Reference number: 154186; MUID:93300536; PMID:8315593
 A:Accession: 154186
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
 A:Molecule type: DNA
 A:Residues: 984-1411 <RE3>
 A:Cross-references: GB:L05642; NID:G181892; PIDN:AAA74506.1; PID:G950344
 A:Accession: 168509
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
 A:Molecule type: DNA
 A:Residues: 1776-1913 <RE2>
 A:Cross-references: GB:L05646; NID:G181896; PIDN:AAA74507.1; PID:G950345
 A:Accession: 168510
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
 A:Molecule type: DNA
 A:Residues: 2850-2979 <ROB>
 A:Cross-references: GB:L05649; NID:G181899; PIDN:AAA74508.1; PID:G950346
 A:Roberts, R.G.; Coffey, A.U.; Bobrow, M.; Bentley, D.R.
 Genomics 13, 942-950, 1992
 A:Title: Determination of the exon structure of the distal portion of the dystrophin gene
 A:Reference number: 154175; MUID:92372062; PMID:1505985
 A:Accession: 154175
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
 A:Molecule type: DNA
 A:Residues: 2980-3685 <RES>
 A:Cross-references: GB:M86903; NID:G181881; PIDN:AAA35779.1; PID:G457519
 A:Benitupel, J.; Hillier, M.; Junkes, B.; Piord, M.; Schwainger, E.; Vosberg, H.P.
 Genomics 10, 551-557, 1991
 A:Title: Analysis of a dystrophin gene deletion by amplification of mRNA isolated from D
 A:Reference number: 154166; MUID:91365360; PMID:1889805
 A:Accession: 154166
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 2250-2254 <REA>
 A:Cross-references: GB:S54699; NID:G235303; PIDN:AA19754.1; PID:G235304
 A:Feener, C.A.; Koenig, M.; Kunkel, L.M.
 Nature 338, 509-511, 1989
 A:Title: Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy
 A:Reference number: S03902; MUID:89181947; PMID:2648158
 A:Accession: S03902
 A:Molecule type: mRNA
 A:Residues: 'MED', 12-32/3377-3408 <PEE>
 A:Cross-references: EMBL:X15148
 C:Comment: Dystrophin is proposed to play a role in anchoring the cytoskeleton to the pl
 C:Comment: Defects in dystrophin are responsible for the Duchenne/Becker muscular dystro
 C:Genetics:
 A:Gene: GDB:DMD
 A:Cross-references: GDB:119850; OMIM:310200
 A:Map position: Xp21.2-Xp21.2
 A:Introns: 11/1, 31/3, 62/3, 88/3, 119/3, 177/2, 217/1, 277/3, 320/3, 383/3, 444/2, 494/
 3, 3055/1, 3075/2, 3096/1, 3121/1, 3188/2, 3217/1, 3269/3, 3325/2, 3362/3, 3408/2, 3421/
 A>Note: the list of introns is incomplete
 C:Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dystro
 C:Keywords: actin binding; alternative splicing; calmodulin binding; cytoskeleton; leuc
 elix
 F:14-233/Domain: alpha-actinin actin-binding domain homology <ACT>
 F:253-327/Region: hinge
 F:338-447/Domain: spectrin/dystrophin repeat homology <SP01>
 F:448-556/Domain: spectrin/dystrophin repeat homology <SP02>
 F:558-667/Domain: spectrin/dystrophin repeat homology <SP03>
 F:668-717/Region: hinge
 F:718-828/Domain: spectrin/dystrophin repeat homology <SP04>
 F:838-934/Domain: spectrin/dystrophin repeat homology <SP05>
 F:938-1045/Domain: spectrin/dystrophin repeat homology <SP06>
 F:1047-1154/Domain: spectrin/dystrophin repeat homology <SP07>
 F:1156-1263/Domain: spectrin/dystrophin repeat homology <SP08>
 F:1265-1367/Domain: spectrin/dystrophin repeat homology <SP09>
 F:1372-1477/Domain: spectrin/dystrophin repeat homology <SP10>
 F:1478-1568/Domain: spectrin/dystrophin repeat homology #status atypical <SP11>
 F:1570-1676/Domain: spectrin/dystrophin repeat homology <SP12>
 F:1678-1782/Domain: spectrin/dystrophin repeat homology <SP13>
 F:1784-1875/Domain: spectrin/dystrophin repeat homology #status atypical <SP14>
 F:1876-1982/Domain: spectrin/dystrophin repeat homology <SP15>

F:1984-2101/Domain: spectrin/dystrophin repeat homology <SP16>
 F:2103-2208/Domain: spectrin/dystrophin repeat homology <SP17>
 F:2210-2316/Domain: spectrin/dystrophin repeat homology <SP18>
 F:2327-2423/Domain: spectrin/dystrophin repeat homology <SP19>
 F:2424-2470/Region: hinge
 F:2471-2577/Domain: spectrin/dystrophin repeat homology <SP20>
 F:2579-2686/Domain: spectrin/dystrophin repeat homology <SP21>
 F:2688-2802/Domain: spectrin/dystrophin repeat homology <SP22>
 F:2804-2931/Domain: spectrin/dystrophin repeat homology <SP23>
 F:2933-3040/Domain: spectrin/dystrophin repeat homology <SP24>
 F:3041-3112/Region: hinge
 F:3055-3092/Domain: WW repeat homology <WW1>
 F:3080-3360/Region: cysteine-rich
 F:3506-3527/Region: leucine zipper motif
 F:3572-3593/Region: leucine zipper motif
 Query Match 5.3%; Score 106; DB 1; Length 3685;
 Best Local Similarity 21.2%; Pred. No. 1.2e+02;
 Matches 70; Conservative 59; Mismatches 121; Indels 80; Gaps 14;
 QY 32 PVPSPAPAPGCTKTVKSKOPLOVTKDGRKRPANDLLITNAVLOTNDLVHIGVKFS 91
 DB 3417 PVDASAPASPQSHDTHSRLEHVASR-LAEMENSGSYLNDISPNSIDDEHLLIQHY 3475
 QY 92 CRFTLREVQERVYALLVPVISKLAQMRQLHPAIAIQSKALFKAEEQLSKVGSST 151
 DB 3476 CQ-----SLNQDSPLSQ-----PSRPAQILISLESE 3501
 QY 152 SQPTLE-TFQDLLHRHPAPAFYLARTAKALQAHWQMLKQVYLLEDVTPPLPKGDVLNFS 210
 DB 3502 ERGELERLADLEEN-----RNLQAEYDRLKQOH--EHKGLSPLPSPPMMPFS 3549
 QY 211 -----DAEDLIDSKLKMREDVLEHEMLVADRRQKREIRQLQELHKQVYVD----- 259
 DB 3550 PQSPRDAR-LIAEAKLRLQHKGRLEARQQLIEDHNK---QLSEQLHRQLQLEPPQAEA 3604
 QY 260 --STGMSPPDPDQNTLAVLRGMVRYLMRSREITVLRGATKQNDVDLSLEGPAKISR 317
 DB 3605 KNGTIVSSPFSIDRSRSSQPMILR-----VGSQTSBSMGEDDL-LSPPDITSG 3655
 QY 318 KQGVTKLNKNGDPFLANEGRRPIYIDGRPV 347
 DB 3656 LEEVMEQLNNS--FPSSGRN---TPGKPM 3680
 RESULT 13
 T28667
 dynein beta heavy chain - Paramesitum tetraurelia
 C:Species: Paramesitum tetraurelia
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Mar-2001
 C:Accession: T28667; PC4340
 R:Kandl, K.A.; Forney, J.D.; Asai, D.J.
 submitted to the EMBL Data Library, January 1995
 A:Description: The dynein genes of Paramesitum: the differential expression of axonemal a
 A:Reference number: Z20502
 A:Accession: T28667
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-4588 <KAN>
 A:Cross-references: EMBL:U19464; NID:G625089; PID:G625090; PIDN:AAA61680.1
 R:Asai, D.J.; Beckwith, S.M.; Kandl, K.A.; Keating, H.H.; Tjandra, H.; Forney, J.D.
 J. Cell Sci. 107, 839-847, 1994
 A:Title: The dynein genes of Paramesitum tetraurelia: Sequences adjacent to the catalytic
 A:Reference number: PC4340; MUID:94334383; PMID:8056840
 A:Accession: PC4340
 A:Molecule type: mRNA
 A:Residues: 1886-2085 <ASA>
 C:Genetics:
 A:Genetic code: SGC5
 A:Introns: 43/2, 113/3, 171/3, 4490/1
 C:Superfamily: dynein heavy chain, ciliary
 C:Keywords: nucleotide binding; P-loop
 F:1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match 5.2%; Score 104.5; DB 2; Length 4588;
Best Local Similarity 21.1%; Pred. No. 2,1e+02;
Matches 59; Conservative 58; Mismatches 72; Indels 91; Gaps 17;

Qy 48 KKSQPLQVTKDGLGRKRPANDLLINAVLQTNDLTSVHLGVKFSGRFTLREVERVALL 107
Db 650 EKQKQPL-LSKD-----ENGLLRVNF-----DPALVRL-----LRV--KYFTLL 686
Qy 108 YDPVTSKLACQAMQRLHEALAIQSKALFSAKEBQLSKVSGTSQPLTFEFDLLHHP 167
Db 687 EQPV-----PESASELYSKN-----DTFRE----- 706

Qy 168 DAFYLARTAKALQAHQWQMLKQVYLLLEDQTVQ-PLPKGPQVL-----NPSDAEDLID 217
Db 707 ---YIVQLEMTVENNFVITQJHPHEPLIKRIEKMDEVLKPGIEHKYKMSNDINKFIE 763

Qy 218 DSKLDMDEVLEHMLVADRROKREIRQLBQELHKW--QVLVDSITGMSPPDQNTLA 275
Db 764 TAKA--TVDEL--HQIV---QMKETLKKIEQALEKFMTKIERKNKFMSPDDYQFLKA 816

Qy 276 VLGRMVRVYLMRSREITLGRATKD--NOIDVDSLSEGPW 313
Db 817 VVQNKL--SIVKNGTSTINKLVKEVLQVKVKKQO--AW 852

RESULT 14
G96522
F11A17.16 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96522
R:theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marzali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96522
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <STO>
A:Cross-references: GB:AE005173; NID:g8778962; PIDN:ADD49768.2; GSPDB:GN00141
C:Genetics:
A:Gene: F11A17.16
A:Map position: 1

Query Match 5.2%; Score 103.5; DB 2; Length 558;
Best Local Similarity 21.5%; Pred. No. 13;
Matches 88; Conservative 51; Mismatches 129; Indels 141; Gaps 18;

Qy 7 VEPGRSGSEPSSS-----EKKVSKASTVPPSPAP-GL 43
Db 214 VESSRLSPSPSPRLPTPLPKFLVPASSLGRDENSFPFAPPPPPPPPPRL 273

Qy 44 TKRVKSKQPLQVTKDGLGRKRPANDL--LINAVALQTNDLTSVHLGVKFSGRFTLREVER 102
Db 274 AKAAQKQSP-----FVSQFLQNLKQNSRNLSQSVNKKQVNSA----- 315

Qy 103 WYALIDPVISKLACQAMQRLHEALIAA-IQSKALFSAKEBQLSKVSGTSQPLTFEFD 161
Db 316 -----HNSIVGEI---QNSAHLIAIKADIERKEFI---NDLIQKVLTTC----- 355

Qy 162 LHHRPDAFYLARTAKALQAHQWQMLKQVYLLLEDQTVQPLPKGDVQLNPSDAEDLID----- 217
Db 356 -----FSDMEDVVKFVDW 368

Qy 218 -DSKLQKWRDE--VLEH---ELMVADRRQK-----REIRQLBQELHKWQVLVDSITGMS 266

Db 369 LDKELATLADRAVLKHPKWEKKADTLQEAAYEYRELKLEKELSSYDDPNHGYVAL 428
Qy 267 PDPNQTLAVLRGMVRYLMRSREITLGRATKDNOIDVDSL-GRPAKISRKGIVTLK 325
Db 429 KQMAN--LDDKSEQRIRLVLRGSSM-RSYQDFKIPVEMWLDGMI CKI--KASITLA 483

Qy 326 NNGDFIANEGRPPIYIDGRVLCGSKWRLSNGSVETASLRPFLLNQ 374
Db 484 KTYMNRVANELQSANRD-----RESTKALLQGVRFAYRTHQ 522

RESULT 15
T42725
actin binding protein ACF7, neural isoform 1 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 03-Nov-2000
C:Accession: T42725
R:Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
Genomics 38, 19-29, 1996
A:Title: Cloning and characterization of mouse ACF7, a novel member of the dyctonin subf.
A:Reference number: Z20900; MUID:97124842; PMID:8954775
A:Accession: T42725
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1825 <BER>
A:Cross-references: EMBL:U67203; NID:g1675221; PID:g1675222; PIDN:ACC52988.1
C:Genetics:
A:Gene: ACF7
A:Map position: 4
C:Superfamily: plectrin; alpha-actinin actin-binding domain homology; ribosomal protein S
C:Keywords: actin binding

Query Match 5.2%; Score 103.5; DB 2; Length 1825;
Best Local Similarity 21.6%; Pred. No. 69;
Matches 71; Conservative 59; Mismatches 124; Indels 75; Gaps 15;

Qy 52 QPLQVTKDGLGRKRPANDLLINAVLQTNDLTSVHLGVKFSGRFTLREVO-----ERMYA 105
Db 540 EEMQKLERAEW--GNDLPSVELQLETOQ--HHTSVK-ELGSSVKEARLYEGKMSQNFH 594

Qy 106 LLYDPVTSKLACQ-----AMRLHPEALAIQSKA-----LFSKAEQQLSKVSGT 151
Db 595 TSYVETLQKLETOYCKLKETSFFMRHQLSHKRVSRRTALILVNGKEBELACD--WSD 653

Qy 152 SQPT-----LETFQDL-----LHHRPDAFYLARTAKALQAHQWL 185
Db 654 SNPNISAKKTYFSELTMELBEGQDVFRSLQDTAEVLSLENHPAKQTVBAYSAVQSQQLQW 713

Qy 186 MKQYLLLEDQTVQPLPKGDVQLNPSDAEDLIDDSKLQKMDREVELEHMLVADRQ--KRE 243
Db 714 MKQCLCVQGHKENVAAVQFQ--FSDARDL--ESFLRLQDSI--KKRYTDRSTLSRL 767

Qy 244 IROLEQLHKQVAVDSITGMSPPDQNTLAVLRGMVRVYLMRSR-----EI 291
Db 768 EDLQDSMDEKEQLQSSSVASLVGRSKTIVQLKPRNPDVHLSTLSVKALCVQRLQEI 827

Qy 292 TLGRATKQNOIDVLSLEGPAWKISRQSG 320
Db 828 TI---CKNDECVLEBDSQRTKMKVISPFG 853

Search completed: March 27, 2003, 05:22:44
Job time : 57 secs

This Page Blank (uspto)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:06:28 ; Search time 23 Seconds
(without alignments)

703.294 Million cell updates/sec

Title: US-09-744-125A-2

Perfect score: 1996
Sequence: 1 NSARGVPEGRCSGSEPPSS.....LINODLIALIRAEAKITPQ 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	5.8	699	1 K122_STRPU	P6872 Strongyloce
2	106.5	5.3	4684	1 PLEI_HUMAN	Q15149 homo sapien
3	106	5.3	3685	1 DMD_HUMAN	P11532 homo sapien
4	103.5	5.2	5327	1 ACF7_MOUSE	O9qxx0 mus musculu
5	103	5.2	810	1 GYRA_BORBU	O51396 borrelia bu
6	102	5.1	3678	1 DMD_MOUSE	P11533 mus musculu
7	101.5	5.1	1109	1 POL_CAEVC	P33459 caprine art
8	101	5.1	3680	1 DMD_CANFA	O97592 canis fam1
9	99.5	5.0	417	1 SYS_MYCSE	P47251 mycoplasma
10	99.5	5.0	496	1 H1A2_ARATH	O06402 arabidopsis
11	99.5	5.0	911	1 T2D1_DROME	P41123 drosophila
12	99.5	5.0	2068	1 P70754_SCHPO	Q09794 echinosacch
13	99	5.0	2244	1 PVR1_SCHPO	P70755 aquilex pyr
14	98.5	4.9	855	1 MUTS_AOUPY	O9nycc3 homo sapien
15	98.5	4.9	4486	1 DYNH_HUMAN	O14203 homo sapien
16	98	4.9	1278	1 DYNH_HUMAN	Q8tuz2 bos taurus
17	98	4.9	1453	1 V373_BOVIN	P75498 mycoplasma
18	97.5	4.9	598	1 LEPB_MYCPN	P33458 escherichia
19	97.5	4.9	1048	1 SBCC_ECOLI	Q9jhu4 mus musculu
20	97.5	4.9	4644	1 DYHC_MOUSE	P38650 rattus norv
21	97.5	4.9	4644	1 DYHC_MOUSE	O90640 gallus gall
22	97	4.9	1225	1 KFA4_CHICK	Q01205 rattus norv
23	96	4.8	442	1 DMD_RAT	P11533 gallus gall
24	96	4.8	3660	1 DMD_CHICK	Q9y264 homo sapien
25	95.5	4.8	503	1 AGP4_HUMAN	O43093 syncephalas
26	95.5	4.8	935	1 KINH_SYNRA	O9upn3 homo sapien
27	95.5	4.8	5430	1 ACF7_HUMAN	P70478 rattus norv
28	95	4.8	870	1 CSX2_SCHPO	P45266 haemophilus
29	95	4.8	2842	1 ABC_FAT	P45266 haemophilus
30	94.5	4.7	458	1 RAD4_HABIN	P45266 haemophilus
31	94.5	4.7	754	1 KATC_ARATH	P53533 synecococc
32	94.5	4.7	883	1 CLPB_SYNP7	P53533 synecococc
33	94.5	4.7	1033	1 Y328_MYCPN	P75310 mycoplasma

34	94.5	4.7	2469	1 TEGU_HSVSA	O01056 herpesvirus
35	94	4.7	986	1 GM13_RAT	O62839 rattus norv
36	94	4.7	1539	1 Y373_HUMAN	O15078 homo sapien
37	93.5	4.7	409	1 ODO2_FUGRU	O90512 fugu rubrip
38	93	4.7	339	1 RLAD_ARCFU	O28761 archaeglob
39	93	4.7	402	1 TAU_CAPHI	O02828 capra hircu
40	93	4.7	404	1 Y349_MYCSE	P47531 mycoplasma
41	93	4.7	447	1 TAU_BOVIN	P29172 bos taurus
42	93	4.7	452	1 Y190_CHLPN	O92842 chlamydia p
43	93	4.7	693	1 APP2_MOUSE	O06335 mus musculu
44	93	4.7	852	1 GCF2_DROME	O9xyp7 drosophila
45	92.5	4.6	578	1 LIPA_MYCPU	O50274 mycoplasma

ALIGNMENTS

```

RESULT 1
K122_STRPU STANDARD; PRT; 699 AA.
ID K122_STRPU
AC P6872;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kinesin-II 85 kDa subunit (KRP-85/95 85 kDa subunit).
GN KRP85.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Euechinozoa; Echinodermata; Echinozoa; Echinozoa;
OC Strongylocentrotus.
ON NCBI_TaxID=7668;
RX MEDLINE=94050179; PubMed=8232586;
RA Scholey J.M.;
RT "Novel heterotrimeric kinesin-related protein purified from sea
RT urchin eggs."
RL Nature 366:268-270(1993).
CC -1- SUBUNIT: Heterotrimer of a 115 kDa subunit (KRP115) and two
CC -1- kinesin-like subunits of 95 kDa (KRP95) and 85 kDa (KRP85).
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L16993; AAA16098.1; -
DR HSSP: P17119; 3KAR.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00225; kinesin_1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KM Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 340 KINESIN_MOTOR (BY SIMILARITY).
FT DOMAIN 341 619 COILED COIL (BY SIMILARITY).
FT DOMAIN 620 699 GLOBULAR (BY SIMILARITY).
FT NP BIND 97 104 ATP (POTENTIAL).
SQ SEQUENCE 699 AA; 78697 MW; 7B386611C808190 CRC64;
Query Match 5.8%; Score 115; DB 1; Length 699;
Best Local Similarity 20.6%; Pred. No. 0.89;
Matches 82; Conservative 58; Mismatches 138; Indels 120; Gaps 16;

```

QY 6 GVEPGRSGSPSSSE-----KKVSKAPSTPPSPAPAPGLTKRVKSKOPLQVT 57
 Db 377 GADDEESGSESSGDEEAGEGCVKKKKKPKRKU--SPETMAAQKKIDEKKALEEK 434
 QY 58 KDLGRWKRPANDLLINAVLOTNDLTSLVHLGVKFSGRFTLREVOERWYALLYPVLSKLAC 117
 Db 435 KD-----MVEDRNTVH-----RELQRR-----ESELHK--A 459
 QY 118 QAMROLHPBALAIAISKA-----LFSKAEOLLSKVSTSTOPTLETFODLLHRHPDAFY 171
 Db 460 QDDQKILNEKNAIOKKILVGVVDLAKSEEQ-----EQLEEQ----- 497
 QY 172 LARTKALQAHQMLKQYVYLEDQYVPLPKGDQVYLFPSDADLI--DDSKLKDMDVYL 229
 Db 498 -----SALBEMKRMKQESWRRKMRMEERBOERMDIEBKYSLSLDPAHAKTKKLKVMYMLM 552
 QY 230 EHELVADRQRQRE-----IRQLEQELHKQVLYVDSITGMSPPDF-----NQT 274
 Db 553 QAKSEVADMQAHQREMEALLENVRELSTLMSMLIDSPFOERQEMIEQYVHMNEDI 612
 QY 275 AVLGRMRYLYMRSEBITLGRATKQNOI--DVDS-----LGGPAWKIRKQGVTKL 324
 Db 613 GEMQJCKVAVYTGNNMRKQTPVADKDKSLAYGADLSNVFLTYNLBGGMKYKPSQG--- 668
 QY 325 KNGGDFPIANEGRPPIYIDGRPVLCGSKWRLSNNSVVE 362
 Db 669 -----KSGRPKTSGRP-KTGKKKQASMASSID 695

RESULT 2

ID_PLE1_HUMAN STANDARD; PRT; 4684 AA.
 AC Q15149; Q16640; Q15148;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1).
 GN PLECL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Placenta;
 RX MEDLINE=96210632; PubMed=8633055;
 RA Liu C.-G., Maercker C., Castanon M.J., Hauptmann R., Wiche G.;
 RT "Human plectin: organization of the gene, sequence analysis, and
 RT chromosome localization (8q24).";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND DISEASE.
 RX MEDLINE=96312447; PubMed=8698233;
 RA McLean W.H.I., Pulkkinen L., Smith F.J.D., Rugg E.L., Lane E.B.,
 RA Bullrich F., Burgess R.E., Amano S., Hudson D.L., Owaribe K.,
 RA McGrath J.A., McMillan J.R., Eady R.A.J., Leigh I.M., Cristiano A.M.,
 RA Uitto J.;
 RT "Loss of plectin causes epidermolysis bullosa with muscular dystrophy:
 RT cDNA cloning and genomic organization.";
 RL Genes Dev. 10:1724-1735(1996).
 RN [3]
 RP VARIANT MD-EBS 1003-GLN--ALA-1005 DEL.
 RX MEDLINE=97049959; PubMed=8894687;
 RA Pulkkinen L., Smith F.J.D., Shimizu H., Murata S., Yacita H.,
 RA Hachisuka H., Nishikawa T., McLean W.H.I., Uitto J.;
 RT "Homozygous deletion mutations in the plectin gene (PLECL1) in patients
 RT with epidermolysis bullosa simplex associated with late-onset
 RT muscular dystrophy.";
 RL Hum. Mol. Genet. 5:1539-1546(1996).
 RN [4]
 RP VARIANT MD-EBS LEU-429 INS.
 RX MEDLINE=21090821; PubMed=1159198;
 RA Bauer J.W., Rouan F., Kofler B., Reniczek G.A., Kornacker I.,

RA Mus W., Hametner R., Klaussegger A., Huber A., Pohla-Gubo G.,
 RA Wiche G., Uitto J., Hantner H.;
 RT "A compound heterozygous one amino-acid insertion/nonsense mutation in
 RT the plectin gene causes epidermolysis bullosa simplex with plectin
 RT deficiency.";
 RL Am. J. Pathol. 158:617-625(2001).
 CC -1 FUNCTION: INTERLINGS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND
 CC MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSES OR
 CC HEMIDESMOSES. COULD ALSO BIND MUSCLE PROTEINS SUCH AS ACTIN TO
 CC MEMBRANE COMPLEXES IN MUSCLE. MAY BE INVOLVED NOT ONLY IN THE
 CC CROSSLINKING AND STABILIZATION OF CYTOSKELETAL INTERMEDIATE
 CC FILAMENTS NETWORK, BUT ALSO IN THE REGULATION OF THEIR DYNAMICS.
 CC -1 SUBUNIT: HOMODIMER OR HOMOTETRAMER.
 CC -1 ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1 TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST LEVELS IN
 CC MUSCLE, HEART, PLACENTA AND SPINAL CORD.
 CC -1 DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH
 CC VIMENTIN, DESMIN, GAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N-
 CC AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.
 CC -1 PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM
 CC INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY).
 CC -1 DISEASE: DEFECTS IN PLECL1 ARE THE CAUSE OF EPIDERMOLYSIS BULLOSA
 CC WITH MUSCULAR DYSTROPHY (MD-EBS OR EB-MD), AN AUTOSOMAL RECESSIVE
 CC DISORDER CHARACTERIZED BY EPIDERMAL BLISTER FORMATION AT THE LEVEL
 CC OF THE HEMIDESMOSE AND ASSOCIATED LATE-ONSET MUSCULAR DYSTROPHY.
 CC -1 SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -1 SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1 SIMILARITY: CONTAINS 3 PLECTIN REPEATS.
 CC -1 SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z54367; CAA91196.1; -;
 DR EMBL; U53204; AAB05427.1; -;
 DR EMBL; U63610; AAB05428.1; -;
 DR EMBL; U63609; AAB05428.1; JOINED.
 DR EMBL; X97053; CAA65765.1; -;
 DR HSSP; Q01082; 1BRK.
 DR Genew; HGNC:9069; PLECL1.
 DR MIM; 601282; -;
 DR MIM; 226670; -;
 DR InterPro; IPR001589; Actbind actinin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR001101; Plectin_repeat.
 DR InterPro; IPR005326; S10_plectin.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00681; Plectin; 19.
 DR Pfam; PF03501; S10_plectin; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00250; PLEC; 33.
 DR SMART; SM00150; SPEC; 5.
 DR PROSITE; PS00019; ACTININ_1; FALSE NEG.
 DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
 DR PROSITE; PS50021; CH; 2.
 DR Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;
 KW Phosphorylation; Alternative splicing; Epidermolysis bullosa;
 KW Disease mutation.
 FT DOMAIN 1 1470 GLOBULAR 1.
 FT 1471 2755 CENTRAL FIBROUS ROD DOMAIN.
 FT DOMAIN 2 2756 4684 GLOBULAR 2.
 FT 175 400 ACTIN-BINDING.
 FT DOMAIN 179 282 CH 1.
 FT DOMAIN 295 397 CH 2.
 FT REPEAT 645 710 SPECTRIN 1.

```

FT REPEAT 740 824 SPECTRIN 2.
FT REPEAT 837 930 SPECTRIN 3.
FT REPEAT 1315 1415 SPECTRIN 4.
FT DOMAIN 1469 2756 COILED COIL (POTENTIAL).
FT REPEAT 2826 2863 PLECTIN 1.
FT REPEAT 2864 2901 PLECTIN 2.
FT REPEAT 2902 2939 PLECTIN 3.
FT REPEAT 2940 2977 PLECTIN 4.
FT REPEAT 2981 3015 PLECTIN 5.
FT REPEAT 3116 3153 PLECTIN 6.
FT REPEAT 3154 3191 PLECTIN 7.
FT REPEAT 3192 3229 PLECTIN 8.
FT REPEAT 3230 3267 PLECTIN 9.
FT REPEAT 3268 3305 PLECTIN 10.
FT REPEAT 3306 3343 PLECTIN 11.
FT REPEAT 3485 3522 PLECTIN 12.
FT REPEAT 3523 3560 PLECTIN 13.
FT REPEAT 3561 3598 PLECTIN 14.
FT REPEAT 3599 3636 PLECTIN 15.
FT REPEAT 3640 3674 PLECTIN 16.
FT REPEAT 3820 3857 PLECTIN 17.
FT REPEAT 3858 3895 PLECTIN 18.
FT REPEAT 3896 3933 PLECTIN 19.
FT REPEAT 3934 3971 PLECTIN 20.
FT REPEAT 3975 4008 PLECTIN 21.
FT REPEAT 4063 4100 PLECTIN 22.
FT REPEAT 4101 4138 PLECTIN 23.
FT REPEAT 4139 4176 PLECTIN 24.
FT REPEAT 4177 4214 PLECTIN 25.
FT REPEAT 4218 4252 PLECTIN 26.
FT REPEAT 4265 4305 PLECTIN 27.
FT REPEAT 4319 4356 PLECTIN 28.
FT REPEAT 4408 4445 PLECTIN 29.
FT REPEAT 4446 4483 PLECTIN 30.
FT REPEAT 4484 4521 PLECTIN 31.
FT REPEAT 4522 4559 PLECTIN 32.
FT REPEAT 4560 4597 PLECTIN 33.
FT DOMAIN 4250 4300 BINDING TO INTERMEDIATE FILAMENTS (BY SIMILARITY).
FT 4 X 4 AA TANDEN REPEATS OF G-S-R-X.
FT MOD_RES 4625 4640 PHOSPHORYLATION (BY CDC2) (BY SIMILARITY).
FT VARSPLIC 1 174 MVAGMLMRDQLRAIYELTFREGVAVAKDRPRSLPHV
FT GVTNLQWRAMASLRAGLVAFETPAMCGFMYLTNEGIAHL
FT ROYLHPREIYASLQRRRVAMVMPARRPRHYQAOGLP
FT GSPPKGGPLTEEQRLTKRELEVSPEITVPATQITLA
FT RPEGPAPAT -> MSGEDAEVAVSEVDSVSGSSGSPSPD
FT TLPNLTGRTORSRRSGGAGSGSVLPDAERAVIRIA (IN
FT ISOFORM 2 AND ISOFORM 3).
FT MISSING (IN ISOFORM 3).
FT L -> LL (IN MD-EBS).
FT /FTID=VAR_011336.
FT VARIANT 409 412 MISSING (IN MD-EBS).
FT VARIANT 429 429 /FTID=VAR_011337.
FT VARIANT 1003 1005 MISSING (IN MD-EBS).
FT CONFLICT 185 185 S -> F (IN REF. 2).
FT CONFLICT 259 259 N -> D (IN REF. 2).
FT CONFLICT 550 550 H -> N (IN REF. 2).
FT CONFLICT 560 560 I -> V (IN REF. 2).
FT CONFLICT 706 706 Q -> R (IN REF. 2).
FT CONFLICT 886 886 N -> Y (IN REF. 2).
FT CONFLICT 1002 1002 V -> A (IN REF. 2).
FT CONFLICT 1309 1309 L -> L (IN REF. 2).
FT CONFLICT 1321 1321 L -> V (IN REF. 2).
FT CONFLICT 1334 1334 V -> L (IN REF. 2).
FT CONFLICT 1534 1534 I -> T (IN REF. 2).
FT CONFLICT 1662 1662 A -> M (IN REF. 2).
FT CONFLICT 1688 1690 WLC -> RLR (IN REF. 2).
FT CONFLICT 1767 1767 Q -> E (IN REF. 2).
FT CONFLICT 1789 1789 L -> A (IN REF. 2).
FT CONFLICT 1910 1910 K -> R (IN REF. 2).
FT CONFLICT 2154 2154 K -> N (IN REF. 2).
FT CONFLICT 2160 2160 S -> R (IN REF. 2).
FT CONFLICT 2215 2215 R -> Q (IN REF. 2).

```

```

FT CONFLICT 2244 2244 S -> A (IN REF. 2).
FT CONFLICT 3027 3027 K -> E (IN REF. 2).
FT CONFLICT 3310 3310 E -> A (IN REF. 2).
FT CONFLICT 3361 3361 F -> L (IN REF. 2).
FT CONFLICT 3408 3408 F -> L (IN REF. 2).
FT CONFLICT 3447 3447 S -> A (IN REF. 2).
FT CONFLICT 3531 3531 G -> A (IN REF. 2).
FT CONFLICT 3580 3580 R -> S (IN REF. 2).
FT CONFLICT 3589 3589 K -> Q (IN REF. 2).
FT CONFLICT 3596 3596 E -> Q (IN REF. 2).
FT CONFLICT 3616 3616 N -> H (IN REF. 2).
FT CONFLICT 3686 3686 V -> A (IN REF. 2).

Query Match 5.3%; Score 106.5; DB 1; Length 4684;
Best Local Similarity 23.5%; Pred. No. 50;
Matches 70; Conservative 52; Mismatches 119; Indels 57; Gaps 13;

Qy 122 QLHPEA-IAAIOKALPFSKAEOLLSKVGSTGPTLETFODLLRHHPDAFYLAATAKALQ 180
Db 2598 KLQGEAKLLQKSEBMTVQGEOLLESTQALQSFLEKSLDR--BRFIEQKAKLEQ 2655

Qy 181 AHW-QLMKQYLLDDQYVQPLPKGDQVLFSDAEDLDDSKLKMREVLNHELMVADR 239
Db 2656 LFQDEVAKAQLREQQRQ-----QQMEQERQRLV--ASMEBARRR--CHEAEQGVRR 2705

Qy 240 QKREIROLQELHKKOVLVDSITGMSPPDNO---TLAVIRGWRVYLRSPREITLGR 295
Db 2706 KQEEIQQLQEQRRQOEELAE-----ENQRLREQQLLEBQCHRAALAHSEVTASQ 2756

Qy 296 --ATKQNOIDVLSLEGPWK-----ISRKQVYIKLKNNGDFPIANEGRRPI---- 340
Db 2757 VAATKTLPLNGRD-ALDGPAAABEPHSPDGLRRKYSAGRLQEG-LSAEELQRLAQGHT 2814

Qy 341 -----YIDGPVTCGSKMRLSNNSVVEISLRFVFLINDLILIRAEAA 385
Db 2815 TVDEIARDEVRYHLQGRSSYIGLLKATNTEKLSVYALQRLSPGTALITLQQA 2872

RESULT 3
DMD_HUMAN STANDARD; PRT; 3685 AA.
ID DMD_HUMAN
AC P11532; Q14169; Q14170;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dystrophin.
GN DMD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=88194521; Pubmed=1282674;
RA Koenig M., Monaco A.P., Kunkel L.M.;
RT "The complete sequence of dystrophin predicts a rod-shaped
RT cytoskeletal protein."
RL Cell 53:219-228(1988).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=89345106; Pubmed=2668885;
RA Rosenthal A., Speer A., Billowitz H., Cross G.S., Forrest S.N.,
RA Davies K.E.;
RT "Two human cDNA molecules coding for the Duchenne muscular dystrophy
RT (DMD) locus are highly homologous."
RL Nucleic Acids Res. 17:5391-5391(1989).
RN [3]
RX SEQUENCE OF 1-497 FROM N.A.
RX MEDLINE=87273512; Pubmed=3607877;
RA Koenig M., Hoffman E.P., Bertelson C.J., Monaco A.P., Feener C.,
RA Kunkel L.M.;
RT "Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
RT preliminary genomic organization of the DMD gene in normal and

```

RT affected individuals." ;
 RL Cell 50:509-517(1987).
 RN [4]
 RP SEQUENCE OF 404-1137 FROM N.A.
 RX MEDLINE=8811512; PubMed=3428261;
 RA Cross G.S., Speer A., Rosenthal A., Forrest S.M., Smith T.J.,
 RA Edwards Y., Flint T., Hill D., Davies K.E.;
 RT "Deletions of fetal and adult muscle cDNA in Duchenne and Becker
 muscular dystrophy patients." ;
 RL EMO J. 6:3277-3283(1987).
 RN [5]
 RP SEQUENCE OF 665-722: 2098-2204 AND 2305-2366 FROM N.A.
 RX MEDLINE=89083552; PubMed=3305741;
 RA Chamberlain J.S., Gibbs R.A., Ranier J.A., Nguyen P.N.,
 RA Caskey C.T.;
 RT "Deletion screening of the Duchenne muscular dystrophy locus via
 multiplex DNA amplification." ;
 RL Nucleic Acids Res. 16:1141-1156(1988).
 RN [6]
 RP SEQUENCE OF 2147-2204 FROM N.A.
 RX MEDLINE=89345155; PubMed=2569720;
 RA Blondel L.A.J., den Dunnen J.T., van Paassen H.M.B.,
 RA Wapenaar M.C., Goetscholtzen P.M., Ginjaar H.B., Bakker E.,
 RA Pearson P.L., van Ommen G.J.B.;
 RT "High resolution deletion breakpoint mapping in the DMD gene by whole
 cosmid hybridization." ;
 RL Nucleic Acids Res. 17:5611-5621(1989).
 RN [7]
 RP SEQUENCE OF 2305-2366 FROM N.A.
 RA Huth A., Will K., Speer A., Bauer D.;
 RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP REVIEW ON DMD POINT MUTATION VARIANTS.
 RX MEDLINE=95038763; PubMed=7951253;
 RA Roberts R.G., Gardner R.J., Bobrow M.;
 RT "Searching for the 1 in 2,400,000: a review of dystrophin gene point
 mutations." ;
 RL Hum. Mutat. 4:1-11(1994).
 RN [9]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=94320940; PubMed=8045556;
 RA Rittinsand F., Reiss J.;
 RT "Microdeletions and polymorphisms in the Duchenne/Becker muscular
 dystrophy gene." ;
 RL Hum. Genet. 94:111-116(1994).
 RN [10]
 RP VARIANT DMD ARG-54.
 RX MEDLINE=94004962; PubMed=8401582;
 RA Prior T.W., Papp A.C., Snyder P.J., Burghes A.H.M., Bartolo C.,
 RA Sedra M.S., Western L.M., Mendell J.R.;
 RT "A missense mutation in the dystrophin gene in a Duchenne muscular
 dystrophy patient." ;
 RL Nat. Genet. 4:357-360(1993).
 RN [11]
 RP VARIANTS DMD HIS-365, TRP-2191 AND ARG-2937.
 RX MEDLINE=95152525; PubMed=7849724;
 RA Nigro V., Nigro G., Esposito M.G., Comi L.I., Molinari A.M.,
 RA Pucca G.A., Pollitano L.;
 RT "Novel small mutations along the DMD/BMD gene associated with
 different phenotypes." ;
 RL Hum. Mol. Genet. 3:1907-1908(1994).
 RN [12]
 RP ALTERNATIVE SPLICING (DYSTROPHIN-1 AND -2).
 RX TISSUE=Retina;
 RA White R.A.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP ALTERNATIVE SPLICING (DYSTROPHIN-3).
 RX TISSUE=Brain;
 RX MEDLINE=89181947; PubMed=2648158;
 RA Feener C.A., Koenig M., Kunkel L.M.;
 RT "Alternative splicing of human dystrophin mRNA generates isoforms at
 the carboxy terminus." ;

RL Nature 338:509-511(1989).
 CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
 CC PLASMA MEMBRANE.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING; THE 3 SHORTER VARIANTS ARE CALLED
 CC DYSTROPHIN-1, -2 AND -3.
 CC -1- DISEASE: Defects in DMD are the cause of Duchenne muscular
 CC dystrophy (DMD) and Becker muscular dystrophy (BMD). DMD is the
 CC most common form of muscular dystrophy; a sex-linked recessive
 CC disorder. It typically presents in boys aged 3 to 7 year as
 CC proximal muscle weakness causing waddling gait, toe-walking,
 CC lordosis, frequent falls, and difficulty in standing up and
 CC climbing up stairs. The pelvic girdle is affected first, then the
 CC shoulder girdle. Progression is steady and most patients are
 CC confined to a wheelchair by age of 10 or 12. Flexion contractures
 CC and scoliosis ultimately occur. About 50% of patients have a lower
 CC IQ than their genetic expectations would suggest. There is no
 CC treatment. BMD resembles DMD in hereditary and clinical features
 CC but is later in onset and more benign.
 CC -1- DISEASE: Defects in DMD are a cause of X-linked dilated
 CC cardiomyopathy (XLCM).
 CC -1- MISCELLANEOUS: THE DMD GENE IS THE LARGEST KNOWN GENE IN HUMANS.
 CC IT IS 2.4 MILLION BASE-PAIRS IN SIZE AND COMPRISES 79 EXONS.
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
 CC ABR-120, ABR-180, OR BETA-FODRIN).
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
 CC -1- DATABASE: NAME=DMD; NOTE=Dystrophin Mutation Database;
 CC WWW="http://www.dmd.nl/database.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X06179; CAA29545.1; ALT_SEQ.
 CC EMBL: X06178; CAA29544.1; -.
 CC EMBL: X14298; CAA32479.1; ALT_SEQ.
 CC EMBL: X15495; CAA33518.1; -.
 CC EMBL: X54820; CAA38589.1; -.
 CC EMBL: X13045; CAA31451.1; -.
 CC EMBL: X13046; CAA31452.1; -.
 CC EMBL: X13047; CAA31453.1; -.
 CC EMBL: X13048; CAA31454.1; -.
 CC EMBL: U27203; AA486115.1; -.
 CC EMBL: U27203; AA486116.1; -.
 CC EMBL: X15148; CAA33245.1; -.
 CC PIR: A27605; A27605.
 CC PIR: A27162; A27162.
 CC PIR: S05291; S05291.
 CC HSSP: P46939; IQAG.
 CC GeneW: HGNC:2928; DMD.
 CC MIM: 300377; -.
 CC MIM: 310200; -.
 CC MIM: 300376; -.
 CC MIM: 302045; -.
 CC InterPro: IPR001589; Actbind_actinin.
 CC InterPro: IPR001715; Calponin-like.
 CC InterPro: IPR002017; Spectrin.
 CC InterPro: IPR002349; WW.
 CC InterPro: IPR001202; WW_Rsp5_WWP.
 CC InterPro: IPR000433; ZnF_ZZ.
 CC Pfam: PF00307; CH; 2.
 CC Pfam: PF00397; WW; 1.
 CC Pfam: PF00435; Spectrin; 22.
 CC Pfam: PF00569; ZZ; 1.
 CC PRINTS: PR00403; WMDOMAIN.

DR SMART; SM00033; CH; 2.
 DR SMART; SM00150; SPEC; 21.
 DR SMART; SM00456; WW; 1.
 DR SMART; SM00291; ZNF_ZZ; 1.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS01159; WW DOMAIN 1; 1.
 DR PROSITE; PS50020; WW DOMAIN 2; 1.
 DR PROSITE; PS01357; ZF_ZZ_1; 1.
 DR PROSITE; PS50135; ZF_ZZ_2; 1.
 DR PROSITE; PS50135; ZF_ZZ_2; 1.
 KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
 KM Repeat; Disease mutation; Actin-binding; Alternative splicing; Zinc-finger.
 FT DOMAIN 1 240 ACTIN-BINDING.
 FT DOMAIN 1 119 CH 1.
 FT DOMAIN 134 237 CH 2.
 FT REPEAT 339 447 SPECTRIN 1.
 FT REPEAT 448 556 SPECTRIN 2.
 FT REPEAT 559 667 SPECTRIN 3.
 FT REPEAT 719 828 SPECTRIN 4.
 FT REPEAT 830 934 SPECTRIN 5.

Query Match 5.3%; Score 106; DB 1; Length 3685;
 Best local similarity 21.2%; Pred. No. 39;
 Matches 70; Conservative 59; Mismatches 121; Indels 80; Gaps 14;

QY 32 PVPSPAPAPGLTKRVKSKOPLOVTKDGRMKRPANDLLINAVLQTNLDSVHGVKFS_91
 DB 3417 PVDAPAPASSPOLSHDTHSRLEHVASR-LAEMENSGSYLNDISPNESIDELHLLIOHY 3475
 QY 92 CRFTLREVEQEWYALVYPVSKLACQAMROLHPEAIAIOSKALFSKAEOLSKVGST 151
 DB 3476 CQ-----SINQSPLSQ-----PSPAQILSLSE 3501
 QY 152 SQPTLE-TEQDLHHRHDPAYLARTAKALQAHWMQKQYLLDQTVPLPKQDVLFNS 210
 DB 3502 ERGELERILADLEEN-----RNLQAEYDRLKQOH-EHKGLSPLSPSPPMPT 3549
 QY 211 -----DAEDLLDSTKCMRDEVLHELMVADRQKREIRQLQELHMKQVLVD----- 259
 DB 3550 PQSPDPAE-LIAEAKLILQHKGRLEARMQILIEDNRK-----QLSQHLRLQLLEQPAEA 3604
 QY 260 --SITGMSSPFDNQTLAVLRGRWYVLMRSREITLGRATKQNOVDVLSLEGPMKISR 317
 DB 3605 KVNQTTVSSTSLQSRSSSQPMILR-----VVGSTISMEEDL-LSPQDITGT 3655
 QY 318 KQGVIKLKNQDFFIANEGREPIYIDGRPV 347
 DB 3656 LEVMEQLNNS--FPSSRGRN---TPGKPM 3680

RESULT 4
 ACP7_MOUSE STANDARD; PRT; 5327 AA.
 ID ACP7_MOUSE
 AC O9QXZ0; P97394; P97395; P97396;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Actin cross-linking family protein 7 (Microtubule actin crosslinking factor) (MACF).
 GN ACP7 OR ACTP7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=BALB/c;
 RX MEDLINE=20068791; PubMed=10601340;
 RA Leung C.U., Sun D., Zheng M., Knowles D.R., Liem R.K.H.;
 RT "Microtubule actin cross-linking factor (MACF): a hybrid of dyctonin
 RT and dystrophin that can interact with the actin and microtubule
 RT cyoskeletons.";

RL J. Cell Biol. 147:1275-1286(1999).
 RN [2]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=97124842; PubMed=8954775;
 RA Bernier G., Mathieu M., De Repentigny Y., Vidal S.M., Kochary R.;
 RT "Cloning and characterization of mouse ACP7, a novel member of the
 RT dyctonin subfamily of actin binding proteins.";
 RL Genomics 38:19-29(1996).
 CC -1- FUNCTION: F-ACTIN-BINDING PROTEIN WHICH MAY PLAY A ROLE IN CROSS-
 CC LINKING ACTIN TO OTHER CYTOSKELETAL PROTEINS. ALSO BINDS TO
 CC MICROTUBULES.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1, 2 (SHOWN HERE) AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN LUNG, BRAIN, SPINAL CORD,
 CC SKELETAL AND CARDIAC MUSCLE, AND SKIN.
 CC -1- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 37 SPECTRIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF150755; AAD32244.1; -;
 DR EMBL; U67203; AAC52988.1; -;
 DR EMBL; U67204; AAC52989.1; -;
 DR EMBL; U67205; AAC52990.1; -;
 DR HSSP; Q01082; 1BKR.
 DR MCD; MG1:108559; Actp7.
 DR InterPro; IPR001589; Actbind actinin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003108; GAS2.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; spectrin; 37.
 DR Pfam; PF02187; GAS2; 1.
 DR ProDom; PD000012; EF-hand; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00054; Efb; 2.
 DR SMART; SM00243; GAS2; 1.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00150; SPEC; 32.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50002; SH3; FALSE_NEG.
 DR PROSITE; PS00018; EF_HAND; 2.
 KW Actin-binding; Cytoskeleton; Calcium-binding; Repeat; SH3 domain;
 KW Alternative splicing.
 FT DOMAIN 1 295 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 78 181 CH 1.
 FT DOMAIN 194 295 SPECTRIN 1.
 FT REPEAT 314 355 SPECTRIN 1.
 FT REPEAT 591 623 SPECTRIN 2.
 FT REPEAT 680 784 SPECTRIN 3.
 FT REPEAT 786 800 SPECTRIN 4.
 FT REPEAT 871 923 SH3.
 FT REPEAT 1250 1272 SPECTRIN 5.
 FT REPEAT 1287 1342 SPECTRIN 6.
 FT REPEAT 1458 1534 SPECTRIN 7.
 FT REPEAT 1593 1660 SPECTRIN 8.
 FT REPEAT 1817 1886 SPECTRIN 9.

FT	REPEAT	1934	2044	SPECTRIN 10.
FT	REPEAT	2262	2282	SPECTRIN 11.
FT	REPEAT	2376	2397	SPECTRIN 12.
FT	REPEAT	2400	2509	SPECTRIN 13.
FT	REPEAT	2557	2617	SPECTRIN 14.
FT	REPEAT	2620	2727	SPECTRIN 15.
FT	REPEAT	2730	2837	SPECTRIN 16.
FT	REPEAT	2840	2944	SPECTRIN 17.
FT	REPEAT	2946	2972	SPECTRIN 18.
FT	REPEAT	2986	3029	SPECTRIN 19.
FT	REPEAT	3086	3162	SPECTRIN 20.
FT	REPEAT	3195	3273	SPECTRIN 21.
FT	REPEAT	3276	3382	SPECTRIN 22.
FT	REPEAT	3385	3491	SPECTRIN 23.
FT	REPEAT	3494	3518	SPECTRIN 24.
FT	REPEAT	3603	3709	SPECTRIN 25.
FT	REPEAT	3722	3817	SPECTRIN 26.
FT	REPEAT	3869	3930	SPECTRIN 27.
FT	REPEAT	3933	4039	SPECTRIN 28.
FT	REPEAT	4042	4149	SPECTRIN 29.
FT	REPEAT	4152	4258	SPECTRIN 30.
FT	REPEAT	4261	4368	SPECTRIN 31.
FT	REPEAT	4371	4477	SPECTRIN 32.
FT	REPEAT	4480	4587	SPECTRIN 33.
FT	REPEAT	4594	4695	SPECTRIN 34.
FT	REPEAT	4698	4804	SPECTRIN 35.
FT	REPEAT	4807	4872	SPECTRIN 36.
FT	REPEAT	4917	4941	SPECTRIN 37.
FT	CA BIND	4987	4998	BE-HAND 1 (POTENTIAL).
FT	CA BIND	5023	5034	BE-HAND 2 (POTENTIAL).
FT	DOMAIN	5173	5180	POLY-SER.
FT	DOMAIN	5252	5287	4 X 4 AA. TANDEM REPEATS OF [GS]-S-R-[AR].
FT	VARSPLIC	1	73	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1).
FT	VARSPLIC	1	181	MISSPELTLSKRSKRSERCSERSYRSRSGSLSPCPGPD
FT	VARSPLIC	1	181	TLPMNLPHCEQKSGSDVLDPAARAVVRA -> EKEFV
FT	VARSPLIC	1	181	QAYEDLERYK (IN ISOFORM 1

```

Db      825 EDLQDSNDEKKEQLIQSSSVASVLGRSKTIVQLKRPBPDHVLKSTLSVKAIQCVQRQLEI 884
Oy      292 TLGRATKQNDIDVDLSLEGPAMKISRKQG 320
Db      885 TI---CKNDECVLEEDNSQRTKMKVISPFG 910

RESULT 5
GTRA_BORBU
ID      GTRA_BORBU STANDARD; PRT; 810 AA.
AC      051396; Q44931;
AD      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      DNA gyrase subunit A (EC 5.99.1.13).
GN      GYRA OR BB0435.
OC      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 35210 / B31;
RX      MEDLINE=98065943; PubMed=9403685;
RA      Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA      Lachigra R., White J.-F., Fleischmann R.D., Richardson D.,
RA      Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA      Peterson J., Kertlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA      van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Meldrum J.,
RA      Uterback T., Matthey L., McDonald L., Artach P., Bowman C.,
RA      Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA      Smith H.O., Venter J.C.;
RT      "Genomic sequence of a Lyme disease spirochete, Borrelia
RT      burgdorferi."
RL      Nature 390:580-586(1997).
RN      [2]
RP      SEQUENCE OF 450-545 FROM N.A.
RC      STRAIN=212;
RX      MEDLINE=93146383; PubMed=1490605;
RT      Old I.G., McDougall J., Saint-Girons I., Davidson B.E.;
RT      "Mapping of genes on the linear chromosome of the bacterium Borrelia
RT      burgdorferi: possible locations for its origin of replication."
RL      FEMS Microbiol. Lett. 78:245-250(1992).
CC      -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC      STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC      INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC      DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC      -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC      of double-stranded DNA.
CC      -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC      BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC      ENZYME FORMS AN A2B2 TETRAMER.
CC      -----
CC      THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC      BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC      THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC      USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC      MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC      ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE001148; AAC66803.1; -.
DR      EMBL; Z12165; CAA78157.1; -.
DR      HSSP; P09097; 1AB4.
DR      TIGR; BB0435; -.
DR      InterPro; IPR002205; DNA_topoisomV.
DR      Pfam; PF00521; DNA_topoisomV, 1.
DR      ProDom; PD000742; DNA_topoisomV, 1.
DR      SMART; SM00434; TOP4C; 1.
DR      TIGRFAMs; TIGR01063; gyra, 1.
KW      Topoisomerase; Isomerase; ATP-binding; Antibiotic resistance;
KW      Complete proteome.
FT      ACT_SITE 124 124 DNA_CLEAVAGE (BY SIMILARITY).

```

FT CONFLICT 469 472 IREE -> YKER (IN REF. 2).
 FT CONFLICT 539 539 F -> Y (IN REF. 2).
 SQ SEQUENCE 810 AA; 91379 MM; 395F3930FEF1650 CRC64;
 Query Match 5.2%; Score 103; DB 1; Length 810;
 Best Local Similarity 21.0%; Pred. No. 7.9;
 Matches 81; Conservative 65; Mismatches 131; Indels 108; Gaps 18;

QY 67 NDLLINAVLQNTDITSHLS-----VKRSCTFREVQERYALLYPVSKLACQ 118
 Db 334 NMLALVNCIPQNLNEELLFEFIEHKNTIERIEFDLRKAKAKA-----VLEGINI- 386
 QY 119 AMRQHPAIAIOSKALPSKAEOLSKVSTSQPLETFODLHRPDAFYLTAKA 178
 Db 387 ALNNI-DEVITIKISSKLAKARELVSNFG-LSFIQANSVLDMLQKLTLEIFKLEE 444
 QY 179 LOAHKQKQYLLLEDQVQPLPKDQVFN-----FSD--AEDLIDSKLD 223
 Db 445 LNILSLIKDY--EDILINPV---RIINIRRETIINLGKFGERRRTKIYDEEVAKT 497
 QY 224 MRDEVLEHVLVADROKREIRLOLEQELHKQVLDVSTGMSPPDQTLAVL----- 277
 Db 498 SMSDLMQKENVVMTTKGFLKLSQNEIKDGTGCK--GLSSFDLNGDEIVIALCVNT 555
 QY 278 -----RGRMYVYLRSR-----ITLG-----RATKDNQIDV 304
 Db 556 HDYLFMISNEGL--YLINAYIKSSRSASKQNSIELINLGDPDEILITKNSKDLTD 613
 QY 305 DLSLEGPWKISR-----KQGVIKLK--NNGDFPIANEGRPPIYIDGRPVLCGSK- 352
 Db 614 YLLLTGASKIARFESTDFKAVKSRGVIVIKLNDKDFVTSAE---IVFKDEKVICLSKK 669
 QY 353 -----WELSNVVEIASLR 367
 Db 670 GSAPFENSRDVLNRTGTQVCGMK 694

RESULT 6
 DMD_MOUSE
 ID_DMD_MOUSE STANDARD; PRT; 3678 AA.
 AC P1153; 060703;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dystrophin.
 GN DMD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92253376; PubMed=1579466;
 RA Bies R.D., Phelps S.F., Cortez M.D., Roberts R., Caskey C.T.,
 RA Chamberlain J.S.;
 RT "Human and murine dystrophin mRNA transcripts are differentially
 expressed during skeletal muscle, heart, and brain development";
 RT Nucleic Acids Res. 20:1725-1731 (1992).
 RN [2]
 RP SEQUENCE OF 1-201 FROM N.A.
 RX MEDLINE=87273512; PubMed=3607877;
 RA Koenig M., Hoffman E.P., Bertelson C.J., Monaco A.P., Feener C.,
 RA Kunzel L.M.;
 RT "Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
 RT preliminary genomic organization of the DMD gene in normal and
 RT affected individuals";
 RT Cell 50:509-517 (1987).
 RN [3]
 RP SEQUENCE OF 300-1390 FROM N.A.
 RX MEDLINE=88018015; PubMed=3659917;
 RA Hoffman E.P., Monaco A.P., Feener C.C., Kunzel L.M.;
 RT "Conservation of the Duchenne muscular dystrophy gene in mice and
 RT humans";

RL Science 238:347-350 (1987).
 RN [4]
 RP SEQUENCE OF 986-1056 FROM N.A.
 RC STRAIN=C57BL/10; TISSUE=Skeletal muscle;
 RX MEDLINE=94154933; PubMed=811539;
 RA Chamberlain J.S., Phelps S.F., Cox G.A., Maichele A.J.,
 RA Greenwood A.D.;
 RT "PCR analysis of muscular dystrophy in mdx mice";
 RT Mol. Cell Biol. Hum. Dis. Ser. 3:167-169 (1993).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RC STRAIN=C57BL/10; TISSUE=Retina;
 RX MEDLINE=95360002; PubMed=7633443;
 RA D'Souza V.N., Nguyen T.M., Morris G.E., Karges W., Pillers D.-A.M.,
 RA Ray P.N.;
 RT "A novel dystrophin isoform is required for normal retinal
 RT electrophysiology";
 RT Hum. Mol. Genet. 4:837-842 (1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
 CC PLASMA MEMBRANE.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 11 DISTINCT ISOFORMS EXIST.
 CC -1- TISSUE SPECIFICITY: DIFFERENTIALLY EXPRESSED DURING SKELETAL
 CC MUSCLE, HEART, AND BRAIN DEVELOPMENT. ALSO EXPRESSED IN
 CC RETINA.
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
 CC ABB-120, ABB-160, OR BETA-FODRIN).
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M68859; AAB02797.1; -
 CC EMBL; M18025; AAA37530.1; -
 CC EMBL; U56724; AAB01216.1; -
 CC EMBL; U15218; AAA87068.1; -
 CC PIR; B27162; B27162.
 CC HSSP; P46939; IQAG.
 CC MGD; MGI:94909; Dmd.
 CC InterPro: IPR001589; Actbind actinin.
 CC InterPro: IPR001715; Calponin-like.
 CC InterPro: IPR002017; Spectrin.
 CC InterPro: IPR002349; WW_Reps_WWP.
 CC InterPro: IPR001202; WW_Domain_2.
 CC Pfam; PF00307; CH; 2.
 CC Pfam; PF00397; WW; 1.
 CC Pfam; PF00435; spectrin; 21.
 CC Pfam; PF00569; ZF; 1.
 CC PRINTS; PR00403; WWDOMAIN.
 CC SMART; SMO0033; CH; 2.
 CC SMART; SMO0150; SPEC; 20.
 CC SMART; SMO0291; ZNF_ZZ; 1.
 CC PROSITE; PS00019; ACTININ_1; 1.
 CC PROSITE; PS00020; ACTININ_2; 1.
 CC PROSITE; PS01159; WW_DOMAIN_1; 1.
 CC PROSITE; PS00020; WW_DOMAIN_2; 1.
 CC PROSITE; PS01357; ZF_ZZ_1; 1.
 CC PROSITE; PS01357; ZF_ZZ_2; 1.
 CC Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
 KW Repeat; Alternative splicing; Zinc-finger.
 FT DOMAIN 1 240 ACTIN-BINDING.
 FT DOMAIN 15 119 CH 1.

```

FT DOMAIN 134 237 CH 2.
FT REPEAT 341 449 SPECTRIN 1.
FT REPEAT 450 558 SPECTRIN 2.
FT REPEAT 561 669 SPECTRIN 3.
FT REPEAT 721 830 SPECTRIN 4.
FT REPEAT 832 936 SPECTRIN 5.
FT REPEAT 945 1047 SPECTRIN 6.
FT REPEAT 1050 1156 SPECTRIN 7.
FT REPEAT 1159 1265 SPECTRIN 8.
FT REPEAT 1268 1369 SPECTRIN 9.
FT REPEAT 1470 1570 SPECTRIN 10.
FT REPEAT 1573 1678 SPECTRIN 11.
FT REPEAT 1681 1782 SPECTRIN 12.
FT REPEAT 1879 1981 SPECTRIN 13.
FT REPEAT 2013 2103 SPECTRIN 14.
FT REPEAT 2106 2210 SPECTRIN 15.
FT REPEAT 2213 2318 SPECTRIN 16.
FT REPEAT 2468 2570 SPECTRIN 17.
FT REPEAT 2573 2679 SPECTRIN 18.
FT REPEAT 2682 2795 SPECTRIN 19.
FT REPEAT 2798 2900 SPECTRIN 20.
FT REPEAT 2902 2924 SPECTRIN 21.
FT REPEAT 2927 3033 SPECTRIN 22.
FT DOMAIN 3048 3081 ZZ-TYPE.
FT ZN FING 3300 3347 D -> H (IN REF. 3).
FT CONFLICT 463 463 S -> F (IN REF. 3).
FT CONFLICT 677 677 V -> L (IN REF. 1; AAB02797).
FT CONFLICT 2337 2337
SQ SEQUENCE 3678 AA; 425810 MW; 1D2E74CFD8035EE CRC64;

Query Match 5.1%; Score 102; DB 1; Length 3678;
Best Local Similarity 20.9%; Pred. No. 75;
Matches 69; Conservative 60; Mismatches 121; Indels 80; Gaps 14;

QY 32 PVPSPAPAGLTTRVYKSKOPLOVTKDGRWKPANDLLINAVIQNDNLVHLGVKES 91
DB 3410 PVDASAPASPGSHDPTDRHIEHYAR-LAEMENSGSYLINDISPNEISIDDEHLILQHY 3468
QY 92 CFTFLREVERWYALLVDVYISKLAQAMKQALPEALAIQSAALFSKAEBOILKVGST 151
DB 3469 CQ-----SLNQDPSLSQ-----PRSPAQILSLSEB 3494
QY 132 SQPTLE-TFQDILHRHPRDAFYLAARTAKALQAHQMLKQYVLLLEDOTVOPLPKDQVLYNS 210
DB 3495 EKGELERRIADLSEEN-----RNIQAEYDRKKQOH-EHKGLSPSPPEMPTIS 3542
QY 211 -----DAEDLIDSKLCKMRDEVLEHELMVADRQKREIROLEQELAKQVILVD----- 259
DB 3543 PQSPRAE-LIAEAKLLRQHKGRLBARMOILBDHNK-----QLESQHLRLRLLEQPOA 3597
QY 260 --STGMSGSPDPDNOQLAVLRGAMVYLMRSREITLGRATKDNQIDVDLSLEGPAKISR 317
DB 3598 KVGNTVSSPSTSLQSDSSQPMILR-----VVGQSISEMGEDL-LSPPDQSTSTG 3648
QY 318 KQGVTKLKNNGDFEIANEGRRPIYIDGRPV 347
DB 3649 LEEVMEQLNNS--FSSSRGN--APGKPV 3673

POL_CAEVC STANDARD; PRT; 1109 AA.
ID POL_CAEVC
AC P33459;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POL. polypeptide (Contains: Protease (Retroposin) (EC 3.1.4.23.-);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.1.26.4)).
GN POL.
OS Caprine arthritis encephalitis virus (strain Cork) (CAEV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11661;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=91021037; PubMed=2171210;
RA Saitarelli M., Querat G., Konings D.A.M., Vigne R., Clements J.E.;
RT "Nucleotide sequence and transcriptional analysis of molecular clones
RT of CAEV which generate infectious virus.";
RL Virology 179:347-364 (1990).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- P1M: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS PROTEIN MAY BE SYNTHESIZED AS A GAG-POL
CC POLYPEPTIDE.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL, M33677; AAA91826.1; ALT_INIT.
CC PIR, B45345; B45345.
CC HSRP, P16088; 1DUT.
CC MEROPS, A02.006; -.
CC InterPro: IPR001995; Asparticase_rtyv.
CC InterPro: IPR001969; Asparticase_site.
CC InterPro: IPR001428; DeoxyUTPase.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR003308; Integrase_Zn.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR004777; RYase.
CC InterPro: IPR001584; Rye.
CC Pfam: PF00077; rtp. 1.
CC Pfam: PF00078; rtp. 1.
CC Pfam: PF00552; Integrase. 1.
CC Pfam: PF00665; rye. 1.
CC Pfam: PF00692; dUTPase. 1.
CC Pfam: PF02022; Integrase_Zn. 1.
CC ProDom: PD000946; DeoxyUTPase. 1.
CC PROSITE: PS00141; ASP_PROTEASE. 1.
CC PROSITE: PS00175; ASP_PROT_RETROV. 1.
CC Hydrolase; Aspartyl protease; RNA-directed DNA polymerase;
CC Transferase; Nuclease; Endonuclease; Polypeptide.
CC CHAIN 1 152 PROTEINASE.
FT CHAIN 153 865 REVERSE TRANSCRIPTASE.
FT CHAIN 866 1109 ENDONUCLEASE.
FT ACT_SITE 68 68 BY SIMILARITY.
SQ SEQUENCE 1109 AA; 127678 MW; 97B2F4B370B03CF3 CRC64;

Query Match 5.1%; Score 101.5; DB 1; Length 1109;
Best Local Similarity 19.7%; Pred. No. 16;
Matches 81; Conservative 54; Mismatches 142; Indels 135; Gaps 20;

QY 21 EKKVSKAP-----STVPSPAPAGLTTRVYKSKOPLOVTKDGRWKPANDLLINAV 75
DB 190 ECKGLGKAPPHWTCTPI-----FCIKKSGKXWMLIDFELNK- 228
QY 76 LOTNDLTSVHL-----GVKFSCTFLREVOERWYAL-LYDVVISKLAQAMKQALPEAI 128
DB 229 -QTEDLTEAQQLPHPGGIQRKXHVITLIDIGAVFTIPVEP-YREYTCFTL--LSPNNL 284
QY 129 AAIOSKALFSKAEBOILSKVGSTSQPTL-ETPQDILHRHPD----- 168
DB 285 G--PCKRYWKVLPQGWKLSPEVYQFTWQELIEDWVIGQHPELQFGIYMDDIYIGSLK 342
QY 169 -----AFYLA-----RTAKALQAHW-----QLMKQYVLLLEDOTVOPLPKGDVYL 207
DB 343 KREIVKDLANYIAQYGTFLPEEKROKGYPAKMLGFEHLHPQTMKQKHTLPBLTKGTITL 402

```

QY 208 NFSADLIDSKLMDREVELEHE-----LMVADRRQREIROLEGLHKQ 255
 DB 403 N-----KIQKVLGELVWRSIGKSIPIPLIKLMEGRDRELOSRKLEHVAHWE 451
 QY 256 VLVDSTGMSPPDPNQTAVIRGMRVRLKRSREITLGRATKNOIDVLSLE--GPW 313
 DB 452 ACRKLEMEG-----NYVNDKDVYQGLAMGDVAIEIYVQEKGP 495
 QY 314 -----KISKQGVILKKN-GDFFIANEGRR-ITIDRPLVCSKMWL 355
 DB 496 UNVHNKLSIPQGVIAQKLTQEVIRTKLIPWLLPKE-----EDWRL 543

RESULT 8
 ID_DMD_CANFA STANDARD; PRT; 3680 AA.
 AC 097592;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dystrophin.
 GN DMD.
 OS Canis familiaris (dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STAIN=Golden retriever;
 RA Cavillie K.S., Mann C.J., Schatzberg S.J., Milton S.D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
 CC PLASMA MEMBRANE.
 CC -1 SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
 CC APP-120, ABP-180, OR BETA-FODRIN).
 CC -1 SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1 SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
 CC -1 SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).

CC -----
 CC EMBL: AF070485; AAC83646.1; --
 DR HSSP; P46939; 10NG.
 DR InterPro: IPR001589; Actbind_actnin.
 DR InterPro: IPR001715; Calponin-like.
 DR InterPro: IPR002017; Spectrin.
 DR InterPro: IPR001202; WW_Rpds_WWP.
 DR InterPro: IPR00433; Znf_ZZ.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00397; WW; 1.
 DR Pfam; PF00435; spectrin; 21.
 DR Pfam; PF00569; Z2; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00150; SPEC; 21.
 DR SMART; SM00456; WW; 1.
 DR SMART; SM00291; Znf_ZZ; 1.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS00021; CH; 2.
 DR PROSITE; PS01159; WW_DOMAIN_1; 1.
 DR PROSITE; PS00020; WW_DOMAIN_2; 1.
 DR PROSITE; PS01357; ZF_ZZ_1; 1.
 DR PROSITE; PS01357; ZF_ZZ_2; 1.
 KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;

KW Repeat; Zinc-finger.
 FT DOMAIN 15 119 CH 1.
 FT DOMAIN 134 237 CH 2.
 FT REPEAT 134 448 SPECTRIN 1.
 FT REPEAT 449 557 SPECTRIN 2.
 FT REPEAT 560 668 SPECTRIN 3.
 FT REPEAT 720 829 SPECTRIN 4.
 FT REPEAT 831 935 SPECTRIN 5.
 FT REPEAT 944 1047 SPECTRIN 6.
 FT REPEAT 1050 1156 SPECTRIN 7.
 FT REPEAT 1159 1265 SPECTRIN 8.
 FT REPEAT 1268 1369 SPECTRIN 9.
 FT REPEAT 1470 1570 SPECTRIN 10.
 FT REPEAT 1573 1678 SPECTRIN 11.
 FT REPEAT 1681 1782 SPECTRIN 12.
 FT REPEAT 1879 1981 SPECTRIN 13.
 FT REPEAT 2013 2103 SPECTRIN 14.
 FT REPEAT 2106 2210 SPECTRIN 15.
 FT REPEAT 2213 2320 SPECTRIN 16.
 FT REPEAT 2470 2572 SPECTRIN 17.
 FT REPEAT 2575 2681 SPECTRIN 18.
 FT REPEAT 2684 2797 SPECTRIN 19.
 FT REPEAT 2800 2902 SPECTRIN 20.
 FT REPEAT 2904 2926 SPECTRIN 21.
 FT REPEAT 2929 3035 SPECTRIN 22.
 FT DOMAIN 3050 3083 WW.
 FT ZN_FING 3302 3349 Z2-TYPE.
 SQ SEQUENCE 3680 AA; 425650 MW; 539F1C9D72377872 CRC64;

Query Match 5.1%; Score 101; DB 1; Length 3680;
 Best Local Similarity 22.8%; Pred. No. 89;
 Matches 66; Conservative 40; Mismatches 83; Indels 101; Gaps 15;

QY 16 EPSSEKKVSKASTPVPSPAPAPGIT-----KRVKSKPLOTYKLGRRK---P 65
 DB 2399 EDLSMDWVTVQLQELERAKPGAPGLTVRAPPSQTVTLVTP-AVTKETAISKLEMP 2457
 QY 66 ANDLLINAVLQTN---DLTSVHLGVFSCRFILREVQERWVALLVDPVI----- 112
 DB 2458 SSLLEVALDLDFRNPATWELTD-----WLSIL-DVILKSQVWYMGD 2497
 QY 113 -----SKLQAMRQLPHEPAIAIOSKALFSKAEQLSKVSGTSQPLETFODL 162
 DB 2498 LEDINEMTIKQKATLQDLQRRPOLBELI-----TAQNLKNTSNGEARITIT--DR 2548
 QY 163 LHRHPDAFYLRKAKALQAHQWLMKQYV-----LLEQOT--VQPLPKGDQVLANSD 211
 DB 2549 IER-----IOSQWDEVQEHQNRRLQUTEMLKDSTOWLEAKKEAEQVLCQAR 2595

QY 212 AEDLIDSKLMDREVELEHELMVADRRQK--EIROLEGLHKQVAVD 259
 DB 2596 A-----KLESWEAP-----YTVDAIQKKTITETKQAKOLRQWQIVND 2633

RESULT 9
 SYS_MYCGE STANDARD; PRT; 417 AA.
 AC P47251;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Seryl-tRNA synthetase (EC 6.1.1.11) (Sertine-cRNA ligase) (Serns).
 GN SERN OR MG005.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 EX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kierlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,

RA Nguyen D.T., Uteerback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium."; *Science* 270:397-403(1995).
 RN [2]
 RN RP SEQUENCE OF 1-125 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94364962; PubMed=8083173;
 RA Bailey C.C., Bott K.F.;
 RT "An unusual gene containing a dnaN N-terminal box flanks the putative
 RT origin of replication of Mycoplasma genitalium."; *J. Bacteriol.* 176:5814-5819(1994).
 RL J. Bacteriol. 176:5814-5819(1994).
 RN [3]
 RN RP SEQUENCE OF 6-112 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III,
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing."; *J. Bacteriol.* 175:7918-7930(1993).
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate
 CC + L-seryl-tRNA(Ser).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39679; AAC1221.1; -
 DR EMBL: U09251; AA57073.1; -
 DR EMBL: U02210; AAD12502.1; -
 DR HSSP: P34945; 1SER.
 DR TIGR: MG005; -
 DR InterPro: IPR002106; AACRNA_ligaseII.
 DR InterPro: IPR002314; tRNA-synt_2b.
 DR InterPro: IPR002317; tRNA-synt_ser.
 DR Pfam: PF00587; tRNA-synt_2b_1.
 DR Pfam: PF02403; Seryl_tRNA_N_1.
 DR PRINTS: PR00981; TRNASYNTHSER.
 DR TIGRFAMs: TIGR00414; sers; 1.
 DR PROSITE: PS50862; AA TRNA_LIGASE II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 KW SEQUENCE 417 AA; 47987 MW; 5E43280A5915DD0 CRC64;
 SQ
 Query Match 5.0%; Score 99.5; DB 1; Length 417;
 Best Local Similarity 23.2%; Pred. No. 57;
 Matches 41; Conservative 35; Mismatches 62; Indels 39; Gaps 8;
 Oy 215 LIDSKLQMDRD-----EVLHEHLMVADRRQREIRFQEHKQGVLDVST 262
 Db 1 MDPNKLNNVDFEKKKLENNVNEQLNQIFQDTKLKRNKQLQDLNANQKSLAKOVA 60
 Oy 263 GMSPPFDNQTALVLRGVRVYLRMSREI-----TLGRATKQNO-IDVLSLEGPAWKIS 316
 Db 61 KKK---DNKKL-----LAESKEKQKTEINNNATKOSNTSODLLNFP--NTA 104
 Oy 317 RKGQVILKNNGDFIFANEGRRPIYIDGRPV-----LCGSKMLSNNSVVEIASLRFV 369
 Db 105 HESVPGKNSANLELLKGRKRVF-DPKPLPHELCEKLNVAFAFKATKISTSTRV 160

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-aminocyclopropane-2-carboxylate synthase 2 (EC 4.4.1.14) (ACC
 DE synthase 2) (S-adenosyl-L-methionine methylthioadenosine-lyase 2).
 GN ACS2 OR ACC1 OR AT1G01480 OR F2214.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eustoidia II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=93028584; PubMed=1357670;
 RA van der Straeten D., Rodriguez-Pousada R.A., Villarreal R., Hanley S.,
 RA Goodman H.M., van Montagu M.;
 RT "Cloning, genetic mapping, and expression analysis of an Arabidopsis
 RT thaliana gene that encodes 1-aminocyclopropane-1-carboxylate
 RT synthase."; *Proc. Natl. Acad. Sci. U.S.A.* 89:9969-9973(1992).
 RL Proc. Natl. Acad. Sci. U.S.A. 89:9969-9973(1992).
 RN [2]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=93066381; PubMed=1438312;
 RA Liang X.-W., Abel S., Keller J.A., Shen N.F., Theologis A.;
 RT "The 1-aminocyclopropane-1-carboxylate synthase gene family of
 RT Arabidopsis thaliana."; *Proc. Natl. Acad. Sci. U.S.A.* 89:11046-11050(1992).
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11046-11050(1992).
 RN [3]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Terry N.;
 RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RL [4]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altabi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultar L.,
 RA Hunter J.U., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.-J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Miltseher J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Uteerback T., Van Aken S., Vayenberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana".
 RL Nature 408:816-820(2000).
 CC -1- FUNCTION: CATALYZES THE FORMATION OF 1-AMINOCYCLOPROPANE-1-
 CC CARBOXYLATE, A DIRECT PRECURSOR OF ETHYLENE IN HIGHER PLANTS.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = 1-
 CC aminocyclopropane-1-carboxylate + methylthioadenosine.
 CC -1- Cofactor: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: Ethylene biosynthesis; first (rate limiting) step.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- TISSUE SPECIFICITY: HIGH IN DEVELOPING LEAVES AND IN FLOWERS.
 CC -1- INDUCTION: BY ETHYLENE.
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sb-sib.ch).

CC -----
 DR EMBL, Z12614, CAA78260.1; -
 DR EMBL, M95595, AAB59298.1; -
 DR EMBL, M95594, AAB97516.1; -
 DR EMBL, Y12776, CAA73310.1; -
 DR EMBL, AC061957, AAF81308.1; -
 DR PIR, S31646, S31646.
 DR PIR, A47199, A47199.
 DR PIR, S46190, S46190.
 DR HSSP, P37821, 188G.
 DR InterPro, IPR001176, ACC synthase.
 DR InterPro, IPR004839, Amino transferase.
 DR InterPro, IPR004838, Nitrates 1.
 DR Pfam, PF00155, aminotran_1.2; 1.
 DR PRINTS, PR00753, ACCSYNTHASE.
 DR PROSITE, PS00105, AA TRANSFER CLASS 1; 1.
 KM Fruit ripening; Ethylene biosynthesis; Lyase; Pyridoxal phosphate;
 KM Multigene family.
 FT BINDING 279
 FT VARIANT 136 136 M -> I.
 SQ SEQUENCE 496 AA; 55531 MW; 766318A9B5F1566 CRC64;

Query March 5.0%; Score 99.5; DB 1; Length 496;
 Best Local Similarity 20.2%; Pred. No. 7.2; Indels 95; Gaps 12;
 Matches 68; Conservative 56; Mismatches 117;

Qy 32 PVPSPAPAPGLTK-----RVKSKQPLQVTKDGRWKPANDLLINAVLOTNDLTSVHL 86
 Db 170 PVPSSSNFPLTVDAAEMAYKKAQSKNKVGLTLTPSNPL---GMLDMDTLTNL-- 224
 Qy 87 GVKSKRTLEAVOERWTALLY---DPVTSKLACQAMRQHPHAIQAQSKALFSKAE 142
 Db 225 -VFVTRKNIHLVVDIEYAAVTFAGDFV-----SVAEVVDVDSIEVNV 268
 Qy 143 QLSAVSGTSGP-----TLTFQDPL---HHPPAFYLAFAKXLAQHWQMKQY 190
 Db 269 DLHIVYLSLQMDGLPGRVGVISFNDVSVCAKMSFGL-----VSSQTLMLASM 322
 Qy 191 LLEDOVQPLPKGDVNLNPSADLIDDSKLDKMRDEVLEHLMVAD----- 237
 Db 323 LSDQFVD-----NF-----LMSSRLGIRHKVFTTGIRKADIACLTSNAGLFAW 368
 Qy 238 --RRQKEIRIQELQELHKQVLDVSTGMSP-----DFDNOTLAVL 277
 Db 369 MDRLHLADRNSFESEIELMHIIIDRVKLVNPSGSSFRCTEPGFRCFANMDDTLHVA 428
 Qy 278 RGRWRYLMRSREITLGRATKDNDIDVLSLEGPAW 313
 Db 429 LGRIDQFVSKNKKIVEKASENDVYONKSAKGLKW 464

RESULT 11
 ID H101 ARATH STANDARD, PRT, 911 AA.
 AC P42730; Q9LE57; Q8W4F2;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heat shock protein 101.
 GN HSP101 OR ATIG74310 OR P1017.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RX MEDLINE=95170291; PubMed=7866032;
 RA Schlomer E.C., Lindquist S., Vierling E.,

RT "An Arabidopsis heat shock protein complements a thermotolerance
 RT defect in yeast."
 RL Plant Cell 6:1899-1909(1994).

RP SEQUENCE FROM N.A.
 RC STRAIN=cv, Columbia;
 RX MEDLINE=20226114; PubMed=10760305;
 RA Hong S.-W., Vierling E.;
 RT "Mutants of Arabidopsis thaliana defective in the acquisition of
 RT tolerance to high temperature stress."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4392-4397(2000).

RP SEQUENCE FROM N.A.
 RC STRAIN=cv, Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federpiet N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Elgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Miltsecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sano H., Salzman J., Schwartz J.R., Shim P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana."
 RL Nature 408:816-820(2000).

RP SEQUENCE OF 1460 FROM N.A.
 RC STRAIN=cv, Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "The Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGSC)."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THERMOTOLERANCE.
 CC -1- INDUCTION: BY HEAT SHOCK.
 CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).

CC -----
 DR EMBL, U13949, AAA67927.1; -
 DR EMBL, AF18796, AAF56423.1; -
 DR EMBL, AC020579, AAG52410.1; -
 DR EMBL, AY062596, AAL32674.1; -
 DR InterPro, IPR003593, AAA_ATPase.
 DR InterPro, IPR003959, AAA_ATPase.
 DR InterPro, IPR001270, Chaprinin_c1pA/B.
 DR InterPro, IPR004176, Clp_N.
 DR Pfam, PF00004, AAA; 2.
 DR Pfam, PF02861, Clp_N; 2.
 DR PRINTS, PR00300, CLP_PTEASEA.
 DR ProDom, PD000739, GSP1_E; 1.
 DR SMART, SM00382, AAA; 2.
 DR PROSITE, PS00870, CLPAB_1; 1.
 DR PROSITE, PS00871, CLPAB_2; 1.
 KM Chaperone; Heat shock; ATP-binding; Repeat.
 FT DOMAIN 164 410
 FT DOMAIN 532 723
 FT NP_BIND 207 214
 FT NP_BIND 606 613
 FT ATP (POTENTIAL).
 FT ATP (POTENTIAL).

```

FT CONFLICT 141 141 V -> F (IN REF. 4).
FT CONFLICT 595 595 P -> A (IN REF. 1).
SQ SEQUENCE 911 AA; 101294 MW; 191EC1853B0C4CB9 CRC64;

Query Match
Best Local Similarity 23.8%; Score 99.5; DB 1; Length 911;
Matches 63; Conservative 42; Mismatches 97; Indels 63; Gaps 11;

QY 105 ALLYPPVSKLACQAMROLHPFAIAIOSKALFSKAE---EQLLSVSGSTSQPLETFPD 161
DB 37 ALISDP-----GIFPQISSGEMNAQSAERVINQALKKLPQSQSPP-----DD 82
QY 162 LHHRPDAFYARTAKALQAHQMLKQVYLLDQVQPLPKGQVQLNFSDAEDLIDDSKL 221
DB 83 I-----PASSSLKIVRRQAQAQKSQSDTLAVDQILM-----GLEDSQI 123
QY 222 KDMRDEVLHEHLMVADROKREIRQLDEQLHKQVYVDSITGMSDFDNOQLAVL---- 277
DB 124 RDLNLEV-----GVAATARKSEVEKLRGKEGK---KVESASG---DTNFQALKTVGRDL 171
QY 278 ---RGRMRYLMRSREI-----TLGRATDNOQIDVDLSLEGPAWKISKQGVYKLNMGD 329
DB 172 VEQAGKLPVIRGRDEIRRVVRLSRITKNIPV---LIGEPGVKTAVVEGLAQRIVKGD 228
QY 330 FFIANGRRPIYIDGRPVLCGSKMR 354
DB 229 VENSITDVRLSLDMGALVAGAKYR 253

RESULT 12
T2D1 DROME STANDARD; PRT; 2068 AA.
AC P51123;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription initiation factor TFIIID 230 kDa subunit (TAFII-230)
DE (TAFII250) (TBP-associated factor 230 kDa) (P230).
GN TAF250.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
OC Muscomorpha; Bphidroides; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-75 AND 540-546.
RX MEDLINE=93279463; PubMed=8504928;
RA Kokubo T., Gong D.-W., Yamashita S., Horikoshi M., Roeder R.G.,
RA Nakatani Y.;
RT "Drosophila 230-kD TFIIID subunit, a functional homolog of the human
RT cell cycle gene product, negatively regulates DNA binding of the TATA
RT box-binding subunit of TFIIID."
RL Genes Dev. 7:1033-1046(1993).
CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE IN TFIIID ASSEMBLY BY
CC INTERACTING WITH BOTH TBP AND OTHER TAF, AS WELL AS SERVING TO
CC LINK THE CONTROL OF TRANSCRIPTION TO THE CELL CYCLE. ESSENTIAL FOR
CC PROGRESSION OF THE G1 PHASE OF THE CELL CYCLE. POSSESSES DNA-
CC BINDING ACTIVITY. IS A NEGATIVE REGULATOR OF THE TATA BOX-BINDING
CC ACTIVITY OF TBP.
CC -!- SUBUNIT: TF2D IS COMPOSED OF TBP AND A VARIETY OF TBP-ASSOCIATED
CC FACTORS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC -!- SIMILARITY: CONTAINS 1 HMG BOX.
CC -!- SIMILARITY: TO HUMAN TAFII-250 (CCG1). SOME TO S.POMBE TAFII-111
CC AND TO S.CEREVISIAE TAF145.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@eb-sib.ch.
CC -----
DR EMBL; S61883; AAB26991.2; -.
DR HSSP; Q92831; 1B91.
DR TRANSFAC; T02119; -.
DR FLYBASE; FBGN010355; TAF250.
DR InterPro; IPR000637; AT hook.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR001878; ZnF CCHC.
DR Pfam; PF00439; Bromodomain; 2.
DR Pfam; PF02178; AT hook; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00384; AT hook; 1.
DR SMART; SM00297; BROMO; 2.
DR SMART; SM00343; ZnF C2HC; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
KW Bromodomain; Nuclear protein; DNA-binding; Cell cycle; Repeat;
KW Transcription regulation; Phosphorylation.
KW DNA BIND 1247 1360 HMG BOX (POTENTIAL).
FT DOMAIN 1445 1451 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 1490 1560 BROMODOMAIN 1.
FT DOMAIN 1612 1682 BROMODOMAIN 2.
FT DOMAIN 1995 2068 GLN-RICH.
FT VARIANT 575 575 P -> S.
SQ SEQUENCE 2068 AA; 232494 MW; AD6A5ABF28B59531 CRC64;

Query Match
Best Local Similarity 21.5%; Pred. No. 51;
Matches 68; Conservative 51; Mismatches 100; Indels 97; Gaps 15;

QY 5 GGVEPGRCS-----GSEPSSEKKVSKAPSTVPVPSPA-----PAPGLTKRKSK 51
DB 1000 GRADPTGCEGFSYRVNPKPQYKEBQSO-PKASVTGTDADRLPLQRAKELLROK 1058
QY 52 OPLQYTKDGRKWPANDLLINAVLOTNDLTSVHLGV-KFS--CRFTLREVERVYALLY 108
DB 1059 VEEBEEKLKSRLV-----EVIDVVRTLSTEKAKGEGMDKFSGNKFSIAHQERY----- 1109
QY 109 DVFISKLACQAMROLHPFAIAIO---SKALFSAEQLSKVG----- 149
DB 1110 -----KEBQRIFFDQNVLLSSEVLSTDEASSSESDLEELKNTLMNLNKKSTQ 1164
QY 150 -STSQPLETFPDLL-----HRHPDAFYARTAKALQAHQML----- 186
DB 1165 LRESEBELR-DELRQLDDEHGGSGGAGKAGKXDPGQQLMATNQGRIILATRTF 1223
QY 187 -----KOYVLLDQVQPLPKGQVQLNFSDAEDLIDSKLKDMRDEVLHEHLMVADRRQ- 240
DB 1224 RNDQKEYTRVE--IVRQPVIDAVI-----KIRTKDQEFIKQFATLDEQOK 1269
QY 241 ---KEIRQLDEQLHK 253
DB 1270 EEMKREKRRIQELRLR 1285

RESULT 13
PRT1 SCHPO STANDARD; PRT; 2244 AA.
AC Q09794;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE URAL protein (includes: Glutamine-dependent carbamoyl phosphate
DE synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2)).
GN URAL OR SPAC22G7.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC NCBI_Taxid=4896;
RN [1]
RP SEQUENCE OF 22-2244 FROM N.A.

```


RC STRAIN=972;
 RX MEDLINE=96020160; PubMed=8590465;
 RA Lollier M., Jaquet L., Nedeva T., Lacroute F., Potier S.,
 RA Souciet J.-L.,
 RT "As in Saccharomyces cerevisiae, aspartate transcarbamoylase is
 assembled on a multifunctional protein including a dihydroorotase-like
 RT cryptic domain in Schizosaccharomyces pombe".
 RL Curr. Genet. 28:138-149 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moul S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakeert G., Aert R., Robben J., Grymopre B.,
 RA Wellens I., Vanstreel E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Beger P., Zimmermann W., Wedler H., Wambutt R., Puntelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lepaute V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt K., Moore K., Hurst S.M.,
 RA Lucas M., Roeder M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipakovski G.V., Useery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe".
 RL Nature 415:871-880 (2002).
 CC -1- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING THREE
 ENZYMATIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPASASE,
 CC AND ATCASE) (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
 CC phosphate + glutamate + carbamoyl phosphate.
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
 CC + N-carbamoyl-L-aspartate.
 CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
 CC -1- PATHWAY: Pyrimidine biosynthesis; second step.
 CC -1- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPASASE
 CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE
 CC GLUTAMINE-DEPENDENT CPASASE (GD-CPASASE) (EC 6.3.5.5).
 CC -1- MISCELLANEOUS: IN EUKARYOTES EC 6.3.5.5 IS SYNTHESIZED BY TWO
 CC PATHWAY-SPECIFIC (ARGININE AND PYRIMIDINE) UNITS SEPARATE CONTROL.
 CC -1- SIMILARITY: THE CPASASE DOMAIN IS SIMILAR TO OTHER CPASASES.
 CC -1- SIMILARITY: IN THE CENTRAL SECTION, BELONGS TO THE DHOASE FAMILY.
 CC DEFECTIVE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X81841; CA57433.1; -;
 DR EMBL; Z54328; CA91130.1; -;
 DR HSPB; P00968; 199X
 DR InterPro; IPR002029; Asp/Om Cotranf.
 DR InterPro; IPR002082; Asp_carbMidtransf.
 DR InterPro; IPR001317; CPS_GATase.
 DR InterPro; IPR000901; CPasase.
 DR InterPro; IPR002474; CPasase_sm_chain.

DR InterPro; IPR002195; Dihydroorotase.
 DR InterPro; IPR000991; GATase_1.
 DR InterPro; IPR004362; MGS_like.
 DR Pfam; PF00117; GATase_1.
 DR Pfam; PF00185; OTCase_1.
 DR Pfam; PF00289; CPasase_L_chain; 2.
 DR Pfam; PF00744; Dihydroorotase; 1.
 DR Pfam; PF00988; CPasase_sm_chain; 1.
 DR Pfam; PF02142; MGS; 1.
 DR Pfam; PF02729; OTCase_N; 1.
 DR Pfam; PF02786; CPasase_L_D2; 2.
 DR Pfam; PF02787; CPasase_L_D3; 1.
 DR PRINTS; PR00100; AOTCase.
 DR PRINTS; PR00098; CPASASE.
 DR PRINTS; PR00099; CPSGATASE.
 DR PRINTS; PR00096; GATASE.
 DR TIGRPFAMS; TIGR00670; asp_caid tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
 DR PROSITE; PS00442; GATASE_TYPE_1; 1.
 DR PROSITE; PS00866; CPASASE_1; 2.
 DR PROSITE; PS00867; CPASASE_2; 2.
 KM Pyrimidine biosynthesis; Ligase; Transferase; Multifunctional enzyme.
 FT DOMAIN 1 437
 FT GATASE (GLUTAMINE AMIDOTRANSFERASE)
 FT (BY SIMILARITY).
 FT LINKER (BY SIMILARITY).
 FT CPASASE (CARBAMOYL-PHOSPHATE SYNTHASE)
 FT (BY SIMILARITY).
 FT LINKER (BY SIMILARITY).
 FT DEFECTIVE DHOASE DOMAIN
 FT (BY SIMILARITY).
 FT LINKER (BY SIMILARITY).
 FT ATCASE (ASPARTATE TRANSAMYLASE)
 FT (BY SIMILARITY).
 FT GATASE (BY SIMILARITY).
 FT GATASE (BY SIMILARITY).
 FT ACT_SITE 422 422
 FT ACT_SITE 424 424
 FT ACT_SITE 424 424
 FT GIC -> RYF (IN REF. 1).
 FT CAVRA -> LQFAQ (IN REF. 1).
 FT EL -> DV (IN REF. 1).
 FT G -> E (IN REF. 1).
 FT G -> E (IN REF. 1).
 FT CONFLICT 1975 1975
 FT CONFLICT 2002 2002
 SQ SEQUENCE 2244 AA; 248306 MW; 5700D15B50CD3B9 CRC64;
 Query Match 5.0%; Score 99; DB 1; Length 2244;
 Best Local Similarity 19.3%; Pred. No. 63; Indels 118; Gaps 19;
 Matches 73; Conservative 65; Mismatches 122;
 QY 15 SEPSSEKKVSKAPSTPPPS-PAPAGLTGRVKSQ-----PLQVTKDLGRWKA 66
 DB 1866 SSASLADSKATRVVSLMSPGLPHAAPSLAFAFGQAPENKAHPDISLMTPN--FKPS 1922
 QY 67 NDLL-LINA-----VLQNDLTSVHLGVFCRFTLRQEQRYALYIPVSKLAC- 117
 DB 1923 HELVOLINSPPYRKHHSVQVTRSDHLVFAAHQRIIVEROGV--DLQYKLLCT 1980
 QY 118 -----QAMROLHPEALIAIQSKALPSKAE-----QLSKVGST---SQPT 155
 DB 1961 MFEPSTRTSSFPAAWQRLGKGVAVTASASVYKGESLADTITRLGCTDAIVLRPS 2040
 QY 156 LFTRODLH-----RHPDAFYL-----ARTA 176
 DB 2041 IESARIANFSPVPLINGNSKEHPQAFDLTYTIRELGSVNGLTITFGDKYGRTV 2100
 QY 177 KALQ---AHQOLMROYLYLEQTYQPLPKGVQ---LNSDAQEDLDSDSKLMQREYV 229
 DB 2101 HSLRLALAFWH-VELHIVSPQLALPDQVKDIPRANGNFIETHEHL-----TKEVVA 2151
 QY 230 EHELMVADRQKREIROLELHKQVYLVDSITGSSPDPNQTIAVLGR--WRYLYMR 287
 DB 2152 QSDVLYCTRVKKEFPASVDE---YKTKDSFT-----VNSVLAASKSCIYVHPLPR 2201
 QY 288 SREITLGRATRONQIDVD 305

Db 2202 NREIS-----EEVDFD 2212

RESULT 14

MUTS_AQUUPY STANDARD; PRT; 855 AA.

AC P70755;

DT 15-JUN-1998 (Rel. 36, Last sequence update)

DT 15-JUN-1998 (Rel. 36, Last sequence update)

DE DNA mismatch repair protein mutS.

GN MUTS.

OS Aquifex pyrophilus.

OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;

CC Aquifex.

CC NCB1_TaxID=2714;

CC NCB1_TaxID=2714;

CC SEQUENCE FROM N.A.

CC Weimur J.G., Rosenfeld A., Wong D.M.;

CC "Hyperthermophilic MutS proteins: isolation, characterization and

CC enhancement of PCR specificity.";

CC Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.

CC -!- FUNCTION: This protein is involved in the repair of mismatches in

CC DNA. It is possible that it carries out the mismatch recognition

CC step. This protein has a weak ATPase activity (By similarity).

CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; U71154; AAB16998.1; -

CC InterPro: IPR000432; Muts_C.

CC InterPro: IPR002863; Muts_N.

CC Pfam; PF00488; Muts_C; 1.

CC Pfam; PF01624; Muts_N; 1.

CC ProDom; PD001263; Muts_C; 1.

CC SMART; SM00534; Muts_C; 1.

CC SMART; SM00533; Muts_N; 1.

CC TIGRfam; TIGR01070; mutS1; 1.

CC PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.

CC DNA repair; ATP-binding; DNA-binding.

CC NP BIND 607 614 ATP (POTENTIAL).

CC SUBMITTED (JUN-1999) to the EMBL/Genbank/DBJ databases.

CC SEQUENCE 855 AA; 97656 MW; 5A8C8EE8E234E8A CRC64;

Query Match 4.9%; Score 98.5; DB 1; Length 855;

Best Local Similarity 21.4%; Pred. No. 18;

Matches 83; Conservative 60; Mismatches 133; Indels 111; Gaps 18;

QY 24 KYSKAPSTVPSPAP--APGLTKRVKSKOPQVTKDL-GWKPANDLLINAVLQTN 80

DB 236 KATQGSFTPLIKRPKYVDEGVKDLKAVKGLTETSEGR---KDLSTLRV---D 287

QY 81 LTVSHLVKVFSCRFLL-----REVQERWYALLVDPVTSIKLACQMRQLHPPEAL 130

DB 288 RLVTGNG--RRRLRFLNPFRSIERIRKQOE-----AVBELINK-----REV 331

QY 131 IOSKALFSAE--EQLISKVSGTSQPTLETFODLIHHPAFIARAKALQAHQWOLMK 189

DB 332 IR-KTLEGMSDERLVRIS--NMASPRELIH-----LKNSLRAEELRKTL 376

QY 190 YLEEDQVQPLPKGDVLFNSDAEDLIDSKLQMDDEVLEHELMV-----ADBRQK 245

DB 377 SLIDSEIFKEIR--GSLNLNKAADLIDKTLVDDPLHVKEGGLIRPGNAYIDEIRFR 434

QY 246 Q-LBOELHMKQVLVDYSITGMS-----PDF--DNQTLAVLNG 279

DB 435 EBAEKLIKVEKKKKETQISQIKGYNKVMGYIEVTYANVYKVEHFRRQTLISNAR 494

QY 280 EMVRYIMRSREITLGRATKNDQIDVDL-----SLEGPAN 313

DB 495 YTTLELQRLERELISAQRIINLELEYELREBEVVKELDKVGNNAATLIGEDVYIQSLAW 554

QY 314 KISRQGVIKLKNNGDFPIANERRP 340

DB 555 IALEKGVKPEVHEGVELIEEGKHVP 581

RESULT 15

DYH9_HUMAN STANDARD; PRT; 4486 AA.

AC Q9NYC9; Q9NO28; Q95494;

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Ciliary dynein heavy chain (Axonemal dynein heavy chain) (Dynein heavy chain 9).

GN DNAH9 OR DNAH17L OR DNAL1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CC NCB1_TaxID=9606;

CC NCB1_TaxID=9606;

CC SEQUENCE FROM N.A.

CC Reed W., Moats-Staats B.M., Carson J.L., Leigh M.W., Collier A.M.;

CC "A ciliary dynein heavy chain whose expression is upregulated in

CC differentiating airway epithelium.";

CC Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.

CC [2]

CC SEQUENCE FROM N.A.

CC Tissue=Nasal epithelium;

CC Bartoloni L., Bloun J.L., Malti A., Sainsbury A., Rossier C.,

CC Gehrig C., She J.X., Maron M.P., Lander E., Meeks M., Chung E.,

CC Jorissen M., Scott H.S., Delozier Blanchet C.D., Gardiner M.,

CC Antonarakis S.E.;

CC "Axonemal beta heavy chain dynein DNAH9: cDNA sequence, genomic

CC structure and investigation of its role in primary ciliary

CC dyskinesia.";

CC Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.

CC [3]

CC SEQUENCE OF 1874-1974 FROM N.A.

CC Tissue=Nasal polyps;

CC Malti A.K., Maltel M.G., Jorissen M., Volz A., Ziegler A.,

CC Bouvagnet P.;

CC "Chromosomal localization of human dynein heavy chain genes.";

CC Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.

CC [4]

CC CHARACTERIZATION.

CC MEDLINE=20556134; PubMed=1104725;

CC Reed W., Carson J.L., Moats-Staats B.M., Lucier T., Hu P.C.,

CC Brighton L., Gambiling T.M., Huang C.H., Leigh M.W., Collier A.M.;

CC "Characterization of an axonemal dynein heavy chain expressed early in

CC airway epithelial cellogenesis.";

CC Am. J. Respir. Cell Mol. Biol. 23:734-741(2000).

CC -!- FUNCTION: FORCE GENERATING PROTEIN OF RESPIRATORY CILIA. PRODUCES

CC FORCE TOWARDS THE MINDS ENDS OF MICROTUBULES. DYNEIN HAS ATPASE

CC ACTIVITY.

CC -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF

CC INTERMEDIATE AND LIGHT CHAINS.

CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; AF257737; AAF69004.1; -

CC EMBL; AJ404468; CAB94756.1; -

CC EMBL; AJ132088; CAA10561.1; -

DR Genew; HGNC:2953; DNAB9.
DR MIM; 603330; -.
DR InterPro; IPR004273; Dynein heavy.
DR Pfam; PF03028; Dynein heavy; 1.
KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT DOMAIN 381 410 COILED COIL (POTENTIAL).
FT DOMAIN 504 529 COILED COIL (POTENTIAL).
FT DOMAIN 639 662 COILED COIL (POTENTIAL).
FT DOMAIN 752 823 COILED COIL (POTENTIAL).
FT DOMAIN 1326 1355 COILED COIL (POTENTIAL).
FT DOMAIN 3051 3154 COILED COIL (POTENTIAL).
FT DOMAIN 3285 3341 COILED COIL (POTENTIAL).
FT DOMAIN 3640 3675 COILED COIL (POTENTIAL).
FT NP_BIND 1870 1877 ATP (POTENTIAL).
FT NP_BIND 2151 2158 ATP (POTENTIAL).
FT NP_BIND 2478 2485 ATP (POTENTIAL).
FT NP_BIND 2825 2832 ATP (POTENTIAL).
FT CONFLICT 2505 2505 L -> V (IN REF. 2).
FT CONFLICT 3678 3678 T -> A (IN REF. 2).
FT CONFLICT 4374 4374 I -> M (IN REF. 2).
SQ SEQUENCE 4486 AA; 511927 MW; 996EDFDEB0B3EB1 CRC64;

Query Match 4.9%; Score 98.5; DB 1; Length 4486;
Best Local Similarity 21.3%; Pred. No. 1.8e+02;
Matches 59; Conservative 47; Mismatches 116; Indels 55; Gaps 11;

QY 127 AIAAIQSKALFSAKEBOLSKVSTGSOPTLET-----FQDLLHRHPD-----AF 170
Db 1372 AVAEIQNPALREIRHWRQDMQATGVSTMDQDTTAAHLQQLHHEDEVRGIVDKAKEM 1431
QY 171 YIARTAKALQAHWQIMKQYLLLEDQTVQPLPKGDQVLNFSDAEDLIDSKLDMRDEVLE 230
Db 1432 GMEKTLKELQTTWAGMERQY-----EPHRTNVPILCSD-EDLI-----EYLE 1473
QY 231 -HELMVADRQKREIRQLQELHAKQVLVDSITGMSPPFDNQL-----AVLRGRWVRY 284
Db 1474 DNQVQLQNLVMSKYVAFLEEVSGWQKLSVDAVISIMFEVQRTWHLESIFTG----- 1528
QY 285 LMRGREITLGRATDNQID-VDSLSEGPAMKSRKQGYIKLKNNGDFIANEGRRPIYID 343
Db 1529 --SEDIRAQLPQDSKRREGIDIDPKELAYDAQIPNVAVQTTNKPGLYEKLED-----IQ 1580
QY 344 GRPVLCGSKWRLSNNSVVEIASLRFVFLINODLLALI 380
Db 1581 GRLCIC-EKALAEYLDTRKLAFPFRYFLSSSDLDIL 1616

Search completed: March 27, 2003, 05:20:21

Job time : 49 secs

This Page Blank (uspto)